Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/019489

International filing date: 18 June 2004 (18.06.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/480,035
Filing date: 19 June 2003 (19.06.2003)

Date of receipt at the International Bureau: 10 September 2004 (10.09.2004)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1219613

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 31, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/480,035*
FILING DATE: *June 19, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/19489*

Certified by

Jon W Dudas

Acting Under Secretary of Commerce
for Intellectual Property
and Acting Director of the U.S.
Patent and Trademark Office

TMH:dv 4239-66168 06/19/03 E-222-2003/0

EXPRESS MAIL LABEL NO. EV339205670US
Date of Deposit: June 19, 2003
PATENT
Attorney Reference Number 4239-66168

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

MAIL STOP PROVISIONAL PATENT APPLICATION
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR § 1.53(c).

TITLE: FUNCTIONAL VARIANTS OF HUMAN TASTE RECEPTOR GENES

Inventor(s)/Applicant(s):

Drayna	**Dennis**		**Potomac, Maryland**
(Last)	(First)	(MI)	(City, State or Foreign Country)
Kim	**Un-Kyung**		**Gaithersburg, Maryland**
(Last)	(First)	(MI)	(City, State or Foreign Country)

Enclosed are:

☒ 87 pages of specification.
☒ 5 sheet(s) of drawings.
☒ 76 pages of Sequence Listing

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ Yes, the name of the U.S. Government agency is the Department of Health and Human Services, National Institutes of Health.

☒ Provisional Filing Fee Amount: ☒ $160, large entity

☒ A check in the amount of $ 160.00 to cover the filing fee is enclosed.

☒ The Commissioner is hereby authorized to charge any additional fees which may be required in connection with the filing of this provisional application and recording any assignment filed herewith, or credit over-payment, to Account No. 02-4550. A copy of this sheet is enclosed.

TMH:dv 4239-66168 06/19/03 E-222-2003/0

EXPRESS MAIL LABEL NO. EV339205670US
Date of Deposit: June 19, 2003
PATENT
Attorney Reference Number 4239-66168

☒ Please return the enclosed postcard to confirm that the items listed above have been received.

Address all telephone calls to Tanya M. Harding, Ph.D. at telephone number (503) 226-7391.

Address all correspondence to Customer Number 36218:

KLARQUIST SPARKMAN, LLP
One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, OR 97204

36218
KLARQUIST SPARKMAN, LLP

Respectfully submitted,

KLARQUIST SPARKMAN, LLP

By _____
Tanya M. Harding, Ph.D.
Registration No. 42,630

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

cc: Docketing

FUNCTIONAL VARIANTS OF HUMAN TASTE RECEPTOR GENES

FIELD

This disclosure relates to the field of taste reception, and more particularly to variations in taste receptors, such as bitter taste receptors including those in the T2R family. It further relates to methods for identifying compounds that interact with taste receptors, including compounds that interact differentially with different variants of a taste receptor.

BACKGROUND

Members of the T2R bitter taste receptor family are G protein-coupled receptor characterized by seven transmembrane domains. In contrast to T1Rs, which also belong to the superfamily of G protein-coupled receptors having a large N-terminal domain, T2R bitter taste receptors generally have a short extracellular N terminus. These cell surface receptors interact with tastants and initiate signaling cascades that culminate in neurotransmitter release.

Individual members of the T2R family exhibit 30%-70% amino acid identity. The most highly conserved sequence motifs reside in the first and last transmembrane segments, and also in the second cytoplasmic loop. The most divergent regions are the extracellular segments, extending partway into the transmembrane helices, possibly reflecting the need to recognize structurally diverse ligands.

Taste sensitivity to the bitter compound phenylthiocarbamide (PTC) and related chemicals is bimodally distributed, and virtually all human populations tested to date contain some people who can (tasters) and some people who cannot taste (nontasters) PTC. The frequency of tasters in North Americans of European ancestry is about 70%. The PTC taste receptor encoded on chromosome 7 was recently identified as a taste receptor that mediates the bitter taste of at least PTC (Kim *et al.*, Science 299:1221-1225, 2003).

Other bitter receptor genes in humans have been identified (Adler *et al.*, *Cell* 100:693, 2000), but currently it is not known which of these genes encode receptors for what other bitter tastants.

SUMMARY

This disclosure provides a comprehensive collection of single nucleotide polymorphisms (SNPs) in bitter taste receptor (T2R) genes. It is believed that a portion of these SNPs define biologically relevant difference between different alleles of the bitter taste receptor genes. Included in the disclosure are sub-sets of the bitter taste receptor SNPs that represent conserved, non-conserved, silent, and truncation mutations in the corresponding proteins, as well as individual allelic sequences for the various bitter taste receptor genes.

The disclosure further provides methods for using the corresponding allelic variants of the taste receptor genes, alone or in various combinations, to test a subject's bitter tasting profile, and to identify and analyze compounds that interact with and/or influence bitter tastes in subjects.

The foregoing and other features and advantages will become more apparent from the following detailed description of several embodiments, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 (including pages 1-5) is a table showing SNPs identified in the indicated T2R bitter taste receptor genes.

SEQUENCE LISTING

The DNA and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand. In the accompanying sequence listing:

SEQ ID NO: 1 shows the coding nucleic acid sequence of bitter taste receptor gene T2R1 (GenBank Accession No. AF227129), and the protein encoded thereby. One SNP is indicated.

SEQ ID NO: 2 shows the protein sequence of the T2R1 bitter taste receptor.

SEQ ID NO: 3 shows the coding nucleic acid sequence of bitter taste receptor gene T2R3 (GenBank Accession No. AF227130), and the protein encoded thereby. Three SNPs are indicated.

SEQ ID NO: 4 shows the protein sequence of the T2R3 bitter taste receptor.

SEQ ID NO: 5 shows the coding nucleic acid sequence of bitter taste receptor gene T2R4 (GenBank Accession No. AF227131), and the protein encoded thereby. Six SNPs are indicated.

SEQ ID NO: 6 shows the protein sequence of the T2R4 bitter taste receptor.

SEQ ID NO: 7 shows the coding nucleic acid sequence of bitter taste receptor gene T2R4 (GenBank Accession No. AF227132), and the protein encoded thereby. Six SNPs are indicated.

SEQ ID NO: 8 shows the protein sequence of the T2R4 bitter taste receptor.

SEQ ID NO: 9 shows the coding nucleic acid sequence of bitter taste receptor gene T2R7 (GenBank Accession No. AF227133), and the protein encoded thereby. One SNP is indicated.

SEQ ID NO: 10 shows the protein sequence of the T2R7 bitter taste receptor.

SEQ ID NO: 11 shows the coding nucleic acid sequence of bitter taste receptor gene T2R8 (GenBank Accession No. AF227134), and the protein encoded thereby. Four SNPs are indicated.

SEQ ID NO: 12 shows the protein sequence of the T2R8 bitter taste receptor.

SEQ ID NO: 13 shows the coding nucleic acid sequence of bitter taste receptor gene T2R9 (GenBank Accession No. AF227135), and the protein encoded thereby. Five SNPs are indicated.

SEQ ID NO: 14 shows the protein sequence of the T2R9 bitter taste receptor.

SEQ ID NO: 15 shows the coding nucleic acid sequence of bitter taste receptor gene T2R10 (GenBank Accession No. AF227136), and the protein encoded thereby. Five SNPs are indicated.

SEQ ID NO: 16 shows the protein sequence of the T2R10 bitter taste receptor.

SEQ ID NO: 17 shows the coding nucleic acid sequence of bitter taste receptor gene T2R13 (GenBank Accession No. AF227137), and the protein encoded thereby. One SNP is indicated.

SEQ ID NO: 18 shows the protein sequence of the T2R13 bitter taste receptor.

SEQ ID NO: 19 shows the coding nucleic acid sequence of bitter taste receptor gene T2R14 (GenBank Accession No. AF227138), and the protein encoded thereby. Two SNPs are indicated.

SEQ ID NO: 20 shows the protein sequence of the T2R14 bitter taste receptor.

SEQ ID NO: 21 shows the coding nucleic acid sequence of bitter taste receptor gene T2R16 (GenBank Accession No. AF227139), and the protein encoded thereby. Seven SNPs are indicated.

SEQ ID NO: 22 shows the protein sequence of the T2R16 bitter taste receptor.

SEQ ID NO: 23 shows the coding nucleic acid sequence of bitter taste receptor gene T2R38 (GenBank Accession No. AF494231), and the protein encoded thereby. Five SNPs are indicated.

SEQ ID NO: 24 shows the protein sequence of the T2R38 bitter taste receptor, also known as the PTC taste receptor.

SEQ ID NO: 25 shows the coding nucleic acid sequence of bitter taste receptor gene T2R39 (GenBank Accession No. AF494230), and the protein encoded thereby. Two SNPs are indicated.

SEQ ID NO: 26 shows the protein sequence of the T2R39 bitter taste receptor.

SEQ ID NO: 27 shows the coding nucleic acid sequence of bitter taste receptor gene T2R40 (GenBank Accession No. AF494229), and the protein encoded thereby. Two SNPs are indicated.

SEQ ID NO: 28 shows the protein sequence of the T2R40 bitter taste receptor.

SEQ ID NO: 29 shows the coding nucleic acid sequence of bitter taste receptor gene T2R41 (GenBank Accession No. AF494232), and the protein encoded thereby. Three SNPs are indicated.

SEQ ID NO: 30 shows the protein sequence of the T2R41 bitter taste receptor.

SEQ ID NO: 31 shows the coding nucleic acid sequence of bitter taste receptor gene T2R43 (GenBank Accession No. AF494237), and the protein encoded thereby. Ten SNPs are indicated.

SEQ ID NO: 32 shows the protein sequence of the T2R43 bitter taste receptor.

SEQ ID NO: 33 shows the coding nucleic acid sequence of bitter taste receptor gene T2R44 (GenBank Accession No. AF494228), and the protein encoded thereby. Ten SNPs are indicated.

SEQ ID NO: 34 shows the protein sequence of the T2R44 bitter taste receptor.

SEQ ID NO: 35 shows the coding nucleic acid sequence of bitter taste receptor gene T2R46 (GenBank Accession No. AF494227), and the protein encoded thereby. Four SNPs are indicated.

SEQ ID NO: 36 shows the protein sequence of the T2R46 bitter taste receptor.

SEQ ID NO: 37 shows the coding nucleic acid sequence of bitter taste receptor gene T2R47 (GenBank Accession No. AF494233), and the protein encoded thereby.

SEQ ID NO: 38 shows the protein sequence of the T2R47 bitter taste receptor.

SEQ ID NO: 39 shows the coding nucleic acid sequence of bitter taste receptor gene T2R44 (GenBank Accession No. AF494234), and the protein encoded thereby. Ten SNPs are indicated.

SEQ ID NO: 40 shows the protein sequence of the T2R44 bitter taste receptor.

SEQ ID NO: 41 shows the coding nucleic acid sequence of bitter taste receptor gene T2R49 (GenBank Accession No. AF494236), and the protein encoded thereby. Ten SNPs are indicated.

SEQ ID NO: 42 shows the protein sequence of the T2R49 bitter taste receptor.

SEQ ID NO: 43 shows the coding nucleic acid sequence of bitter taste receptor gene T2R50 (GenBank Accession No. AF494235), and the protein encoded thereby.

SEQ ID NO: 44 shows the protein sequence of the T2R50 bitter taste receptor.

SEQ ID NO: 45 shows the coding nucleic acid sequence of bitter taste receptor gene T2R60 (GenBank Accession No. AY114094), and the protein encoded thereby. Two SNPs are indicated.

SEQ ID NO: 46 shows the protein sequence of the T2R60 bitter taste receptor.

DETAILED DESCRIPTION

I. Abbreviations

2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
ASO	allele-specific oligonucleotide
ASOH	allele-specific oligonucleotide hybridization
DASH	dynamic allele-specific hybridization
ELISA	enzyme-linked immunosorbant assay
HPLC	high pressure liquid chromatography
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
PCR	polymerase chain reaction
RT-PCR	reverse-transcription polymerase chain reaction
SNP	single nucleotide polymorphism
SSCP	single-strand conformation polymorphism

II. Terms

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes V*, published by Oxford University Press, 1994 (ISBN 0-19-

854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

In order to facilitate review of the various embodiments of the invention, the following explanations of specific terms are provided:

Addressable: Capable of being reliably and consistently located and identified, as in an addressable location on an array.

Amplified RNA (amRNA): A molecule of RNA generated through *in vitro* transcription with T7 or other promoter region attached to the 5' end of the template.

Antisense, Sense, and Antigene: Double-stranded DNA (dsDNA) has two strands, a 5' -> 3' strand, referred to as the plus strand, and a 3' -> 5' strand (the reverse complement), referred to as the minus strand. Because RNA polymerase adds nucleic acids in a 5' -> 3' direction, the minus strand of the DNA serves as the template for the RNA during transcription. Thus, the RNA formed will have a sequence complementary to the minus strand and identical to the plus strand (except that U is substituted for T).

Antisense molecules are molecules that are specifically hybridizable or specifically complementary to either RNA or the plus strand of DNA. Sense molecules are molecules that are specifically hybridizable or specifically complementary to the minus strand of DNA. Antigene molecules are either antisense or sense molecules directed to a dsDNA target.

Array: An arrangement of molecules, particularly biological macromolecules (such as polypeptides or nucleic acids) or biological samples (such as tissue sections) in addressable locations on a substrate, usually a flat substrate such as a membrane, plate or slide. The array may be regular (arranged in uniform rows and columns, for instance) or irregular. The number of addressable locations on the array can vary, for example from a few (such as three) to more than 50, 100, 200, 500, 1000, 10,000, or more. A "microarray" is an array that is miniaturized to such an extent that it benefits from microscopic examination for evaluation.

Within an array, each arrayed molecule (*e.g.*, oligonucleotide) or sample (more generally, a "feature" of the array) is addressable, in that its location can be reliably and consistently determined within the at least two dimensions on the array surface. Thus, in ordered arrays the location of each feature is usually assigned to a sample at the time when it is spotted onto or otherwise applied to the array surface, and a key may be provided in order to correlate each location with the appropriate feature.

Often, ordered arrays are arranged in a symmetrical grid pattern, but samples could be arranged in other patterns (*e.g.*, in radially distributed lines, spiral lines, or ordered clusters). Arrays are computer readable, in that a computer can be programmed to correlate a particular address on the array with information (such as identification of the arrayed sample and hybridization or binding data, including for instance signal intensity). In some examples of computer readable array formats, the individual spots on the array surface will be arranged regularly, for instance in a Cartesian grid pattern, that can be correlated to address information by a computer.

The sample application spot (or feature) on an array may assume many different shapes. Thus, though the term "spot" is used herein, it refers generally to a localized deposit of nucleic acid or other biomolecule, and is not limited to a round or substantially round region. For instance, substantially square regions of application can be used with arrays, as can be regions that are substantially rectangular (such as a slot blot-type application), or triangular, oval, irregular, and so forth. The shape of the array substrate itself is also immaterial, though it is usually substantially flat and may be rectangular or square in general shape.

Binding or interaction: An association between two substances or molecules, such as the hybridization of one nucleic acid molecule to another (or itself). Disclosed arrays are used to detect binding of, in some embodiments, a labeled nucleic acid molecule (target) to an immobilized nucleic acid molecule (probe) in one or more features of the array. A labeled target molecule "binds" to a nucleic acid molecule in a spot on an array if, after incubation of the (labeled) target molecule (usually in solution or suspension) with or on the array for a period of time (usually 5 minutes or more, for instance 10 minutes, 20 minutes, 30 minutes, 60 minutes, 90 minutes, 120 minutes or more, for instance over night or even 24 hours),

a detectable amount of that molecule associates with a nucleic acid feature of the array to such an extent that it is not removed by being washed with a relatively low stringency buffer (*e.g.*, higher salt (such as 3 x SSC or higher), room temperature washes). Washing can be carried out, for instance, at room temperature, but other temperatures (either higher or lower) also can be used. Targets will bind probe nucleic acid molecules within different features on the array to different extents, based at least on sequence homology, and the term "bind" encompasses both relatively weak and relatively strong interactions. Thus, some binding will persist after the array is washed in a more stringent buffer (*e.g.*, lower salt (such as about 0.5 to about 1.5 x SSC), 55-65° C washes).

Where the probe and target molecules are both nucleic acids, binding of the test or reference molecule to a feature on the array can be discussed in terms of the specific complementarity between the probe and the target nucleic acids. Also contemplated herein are protein-based arrays, where the probe molecules are or comprise proteins, and/or where the target molecules are or comprise proteins.

cDNA: A DNA molecule lacking internal, non-coding segments (*e.g.*, introns) and regulatory sequences that determine transcription. By way of example, cDNA may be synthesized in the laboratory by reverse transcription from messenger RNA extracted from cells.

DNA (deoxyribonucleic acid): DNA is a long chain polymer that contains the genetic material of most living organisms (the genes of some viruses are made of ribonucleic acid (RNA)). The repeating units in DNA polymers are four different nucleotides, each of which includes one of the four bases (adenine, guanine, cytosine and thymine) bound to a deoxyribose sugar to which a phosphate group is attached. Triplets of nucleotides (referred to as codons) code for each amino acid in a polypeptide, or for a stop signal. The term "codon" is also used for the corresponding (and complementary) sequences of three nucleotides in the mRNA into which the DNA sequence is transcribed.

Enriched: The term "enriched" means that the concentration of a material is at least about 2, 5, 10, 100, or 1000 times its natural concentration (for example), advantageously at least 0.01% by weight. Enriched preparations of about 0.5%, 1%, 5%, 10%, and 20% by weight are also contemplated.

EST (Expressed Sequence Tag): A partial DNA or cDNA sequence, typically of between 200 and 2000 sequential nucleotides, obtained from a genomic or cDNA library, prepared from a selected cell, cell type, tissue or tissue type, organ or organism, which corresponds to an mRNA of a gene found in that library. An EST is generally a DNA molecule sequenced from and shorter than the cDNA from which it is obtained.

Fluorophore: A chemical compound, which when excited by exposure to a particular wavelength of light, emits light (*i.e.*, fluoresces), for example at a different wavelength. Fluorophores can be described in terms of their emission profile, or "color." Green fluorophores, for example Cy3, FITC, and Oregon Green, are characterized by their emission at wavelengths generally in the range of 515-540 λ. Red fluorophores, for example Texas Red, Cy5 and tetramethylrhodamine, are characterized by their emission at wavelengths generally in the range of 590-690 λ.

Examples of fluorophores that may be used are provided in U.S. Patent No. 5,866,366 to Nazarenko *et al.*, and include for instance: 4-acetamido-4'-isothiocyanatostilbene-2,2'disulfonic acid, acridine and derivatives such as acridine and acridine isothiocyanate, 5-(2'-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS), 4-amino-N-[3-vinylsulfonyl)phenyl]naphthalimide-3,5 disulfonate (Lucifer Yellow VS), N-(4-anilino-1-naphthyl)maleimide, anthranilamide, Brilliant Yellow, coumarin and derivatives such as coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcouluarin (Coumaran 151); cyanosine; 4',6-diaminidino-2-phenylindole (DAPI); 5', 5"-dibromopyrogallol-sulfonephthalein (Bromopyrogallol Red); 7-diethylamino-3-(4'-isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4'-diisothiocyanatodihydro-stilbene-2,2'-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid; 5-[dimethylamino]naphthalene-1-sulfonyl chloride (DNS, dansyl chloride); 4-(4'-dimethylaminophenylazo)benzoic acid (DABCYL); 4-dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives such as eosin and eosin isothiocyanate; erythrosin and derivatives such as erythrosin B and erythrosin isothiocyanate; ethidium; fluorescein and derivatives such as 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-yl)aminofluorescein (DTAF), 2'7'-dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE), fluorescein, fluorescein

isothiocyanate (FITC), and QFITC (XRITC); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferone; ortho cresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthaldialdehyde; pyrene and derivatives such as pyrene, pyrene butyrate and succinimidyl 1-pyrene butyrate; Reactive Red 4 (Cibacron .RTM. Brilliant Red 3B-A); rhodamine and derivatives such as 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B sulfonyl chloride, rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B, sulforhodamine 101 and sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); tetramethyl rhodamine; tetramethyl rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid and terbium chelate derivatives.

Other contemplated fluorophores include GFP (green fluorescent protein), Lissamine™, diethylaminocoumarin, fluorescein chlorotriazinyl, naphthofluorescein, 4,7-dichlororhodamine and xanthene and derivatives thereof. Other fluorophores known to those skilled in the art may also be used.

High throughput genomics: Application of genomic or genetic data or analysis techniques that use microarrays or other genomic technologies to rapidly identify large numbers of genes or proteins, or distinguish their structure, expression or function from normal or abnormal cells or tissues, or from cells or tissues of subjects with known or unknown phenotype and/or genotype.

Human Cells: Cells obtained from a member of the species *Homo sapiens*. The cells can be obtained from any source, for example peripheral blood, urine, saliva, tissue biopsy, surgical specimen, amniocentesis samples and autopsy material. From these cells, genomic DNA, mRNA, cDNA, RNA, and/or protein can be isolated.

Hybridization: Nucleic acid molecules that are complementary to each other hybridize by hydrogen bonding, which includes Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding between complementary nucleotide units. For example, adenine and thymine are complementary nucleobases that pair through formation of hydrogen bonds. "Complementary" refers to sequence complementarity between two nucleotide units. For example, if a nucleotide unit at

a certain position of an oligonucleotide is capable of hydrogen bonding with a nucleotide unit at the same position of a DNA or RNA molecule, then the oligonucleotides are complementary to each other at that position. The oligonucleotide and the DNA or RNA are complementary to each other when a sufficient number of corresponding positions in each molecule are occupied by nucleotide units which can hydrogen bond with each other.

"Specifically hybridizable" and "complementary" are terms that indicate a sufficient degree of complementarity such that stable and specific binding occurs between the oligonucleotide and the DNA or RNA or PNA target. An oligonucleotide need not be 100% complementary to its target nucleic acid sequence to be specifically hybridizable. An oligonucleotide is specifically hybridizable when binding of the oligonucleotide to the target DNA or RNA molecule interferes with the normal function of the target DNA or RNA, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligonucleotide to non-target sequences under conditions in which specific binding is desired, for example under physiological conditions in the case of *in vivo* assays, or under conditions in which the assays are performed.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing DNA used. Generally, the temperature of hybridization and the ionic strength (especially the Na^+ concentration) of the hybridization buffer will determine the stringency of hybridization. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed by Sambrook *et al.* in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989), chapters 9 and 11, herein incorporated by reference.

***In vitro* amplification:** Techniques that increase the number of copies of a nucleic acid molecule in a sample or specimen. An example of *in vitro* amplification is the polymerase chain reaction, in which a biological sample collected from a subject is contacted with a pair of oligonucleotide primers, under conditions that allow for the hybridization of the primers to nucleic acid template in the sample. The primers are extended under suitable conditions, dissociated from

the template, and then re-annealed, extended, and dissociated to amplify the number of copies of the nucleic acid.

The product of *in vitro* amplification may be characterized by electrophoresis, restriction endonuclease cleavage patterns, oligonucleotide hybridization or ligation, and/or nucleic acid sequencing, using standard techniques.

Other examples of *in vitro* amplification techniques include strand displacement amplification (see U.S. Patent No. 5,744,311); transcription-free isothermal amplification (see U.S. Patent No. 6,033,881); repair chain reaction amplification (see WO 90/01069); ligase chain reaction amplification (see EP-A-320 308); gap filling ligase chain reaction amplification (see U.S. Patent No. 5,427,930); coupled ligase detection and PCR (see U.S. Patent No. 6,027,889); and NASBA™ RNA transcription-free amplification (see U.S. Patent No. 6,025,134).

Isolated: An "isolated" biological component (such as a nucleic acid molecule, protein or organelle) has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, *i.e.*, other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

Label: Detectable marker or reporter molecules, which can be attached to nucleic acids. Typical labels include fluorophores, radioactive isotopes, ligands, chemiluminescent agents, metal sols and colloids, and enzymes. Methods for labeling and guidance in the choice of labels useful for various purposes are discussed, *e.g.*, in Sambrook *et al.*, in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989) and Ausubel *et al.*, in *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley-Intersciences (1987).

Mutation: Any change of the DNA sequence within a gene or chromosome. In some instances, a mutation will alter a characteristic or trait (phenotype), but this is not always the case. Types of mutations include base substitution point mutations

(*e.g.*, transitions or transversions), deletions, and insertions. Missense mutations are those that introduce a different amino acid into the sequence of the encoded protein; nonsense mutations are those that introduce a new stop codon. In the case of insertions or deletions, mutations can be in-frame (not changing the frame of the overall sequence) or frame shift mutations, which may result in the misreading of a large number of codons (and often leads to abnormal termination of the encoded product due to the presence of a stop codon in the alternative frame).

This term specifically encompasses variations that arise through somatic mutation, for instance those that are found only in disease cells, but not constitutionally, in a given individual. Examples of such somatically-acquired variations include the point mutations that frequently result in altered function of various genes that are involved in development of cancers. This term also encompasses DNA alterations that are present constitutionally, that alter the function of the encoded protein in a readily demonstrable manner, and that can be inherited by the children of an affected individual. In this respect, the term overlaps with "polymorphism," as defined below, but generally refers to the subset of constitutional alterations.

Nucleic acid: A deoxyribonucleotide or ribonucleotide polymer in either single or double stranded form, and unless otherwise limited, encompassing known analogues of natural nucleotides that hybridize to nucleic acids in a manner similar to naturally occurring nucleotides.

Nucleic acid array: An arrangement of nucleic acids (such as DNA or RNA) in assigned locations on a matrix, such as that found in cDNA arrays, or oligonucleotide arrays.

Nucleic acid molecules representing genes: Any nucleic acid, for example DNA (intron or exon or both), cDNA or RNA, of any length suitable for use as a probe or other indicator molecule, and that is informative about the corresponding gene.

Nucleotide: "Nucleotide" includes, but is not limited to, a monomer that includes a base linked to a sugar, such as a pyrimidine, purine or synthetic analogs thereof, or a base linked to an amino acid, as in a peptide nucleic acid (PNA). A

nucleotide is one monomer in a polynucleotide. A nucleotide sequence refers to the sequence of bases in a polynucleotide.

Oligonucleotide: A linear single-stranded polynucleotide sequence ranging in length from 2 to about 5,000 bases, for example a polynucleotide (such as DNA or RNA) which is at least 6 nucleotides, for example at least 10, 12, 15, 18, 20, 25, 50, 100, 200, 1,000, or even 5,000 nucleotides long. Oligonucleotides are often synthetic but can also be produced from naturally occurring polynucleotides.

An oligonucleotide analog refers to moieties that function similarly to oligonucleotides but have non-naturally occurring portions. For example, oligonucleotide analogs can contain non-naturally occurring portions, such as altered sugar moieties or inter-sugar linkages, such as a phosphorothioate oligodeoxynucleotide. Functional analogs of naturally occurring polynucleotides can bind to RNA or DNA, and include peptide nucleic acid (PNA) molecules. Such analog molecules may also bind to or interact with polypeptides or proteins.

Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

Open reading frame (ORF): A series of nucleotide triplets (codons) coding for amino acids without any internal termination codons. These sequences are usually translatable into a peptide.

Peptide Nucleic Acid (PNA): An oligonucleotide analog with a backbone comprised of monomers coupled by amide (peptide) bonds, such as amino acid monomers joined by peptide bonds.

Pharmaceutically acceptable carriers: The pharmaceutically acceptable carriers useful with compositions provided herein are conventional. By way of example, Martin, in *Remington's Pharmaceutical Sciences*, published by Mack Publishing Co., Easton, PA, 19th Edition, 1995, describes compositions and

formulations suitable for pharmaceutical delivery of the nucleotides and proteins herein disclosed.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (*e.g.*, powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

Polymorphism: Variant in a sequence of a gene, usually carried from one generation to another in a population. Polymorphisms can be those variations (nucleotide sequence differences) that, while having a different nucleotide sequence, produce functionally equivalent gene products, such as those variations generally found between individuals, different ethnic groups, geographic locations. The term polymorphism also encompasses variations that produce gene products with altered function, *i.e.*, variants in the gene sequence that lead to gene products that are not functionally equivalent. This term also encompasses variations that produce no gene product, an inactive gene product, or increased or increased activity gene product.

Polymorphisms can be referred to, for instance, by the nucleotide position at which the variation exists, by the change in amino acid sequence caused by the nucleotide variation, or by a change in some other characteristic of the nucleic acid molecule or protein that is linked to the variation (*e.g.*, an alteration of a secondary structure such as a stem-loop, or an alteration of the binding affinity of the nucleic acid for associated molecules, such as polymerases, RNases, and so forth).

Probes and primers: Nucleic acid probes and primers can be readily prepared based on the nucleic acid molecules provided as indicators of taste reception or likely taste reception. It is also appropriate to generate probes and primers based on fragments or portions of these nucleic acid molecules, particularly

in order to distinguish between and among different alleles and haplotypes within a single gene. Also appropriate are probes and primers specific for the reverse complement of these sequences, as well as probes and primers to 5' or 3' regions.

A probe comprises an isolated nucleic acid attached to a detectable label or other reporter molecule. Typical labels include radioactive isotopes, enzyme substrates, co-factors, ligands, chemiluminescent or fluorescent agents, haptens, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, *e.g.*, in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989) and Ausubel *et al.* (In *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998).

Primers are short nucleic acid molecules, for instance DNA oligonucleotides 10 nucleotides or more in length. Longer DNA oligonucleotides may be about 15, 20, 25, 30 or 50 nucleotides or more in length. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then the primer extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, *e.g.*, by the polymerase chain reaction (PCR) or other *in vitro* nucleic-acid amplification methods known in the art.

Methods for preparing and using nucleic acid probes and primers are described, for example, in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989), Ausubel *et al.* (ed.) (In *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998), and Innis *et al.* (*PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990). Amplification primer pairs (for instance, for use with polymerase chain reaction amplification) can be derived from a known sequence such as any of the bitter taste receptor sequences and specific alleles thereof described herein, for example, by using computer programs intended for that purpose such as PRIMER (Version 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA).

One of ordinary skill in the art will appreciate that the specificity of a particular probe or primer increases with its length. Thus, for example, a primer comprising 30 consecutive nucleotides of a bitter taste receptor protein encoding

nucleotide will anneal to a target sequence, such as homolog of a designated taste receptor protein, with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers can be selected that comprise at least 20, 23, 25, 30, 35, 40, 45, 50 or more consecutive nucleotides of a taste receptor gene.

Also provided are isolated nucleic acid molecules that comprise specified lengths of bitter taste receptor-encoding nucleotide sequences. Such molecules may comprise at least 10, 15, 20, 23, 25, 30, 35, 40, 45 or 50 or more (*e.g.*, at least 100, 150, 200, 250, 300 and so forth) consecutive nucleotides of these sequences or more. These molecules may be obtained from any region of the disclosed sequences (*e.g.*, a specified nucleic acid may be apportioned into halves or quarters based on sequence length, and isolated nucleic acid molecules may be derived from the first or second halves of the molecules, or any of the four quarters, etc.). A cDNA or other encoding sequence also can be divided into smaller regions, *e.g.* about eighths, sixteenths, twentieths, fiftieths, and so forth, with similar effect.

Another mode of division, provided by way of example, is to divide a bitter taste receptor sequence based on the regions of the sequence that are relatively more or less homologous to other bitter taste receptor sequences.

Nucleic acid molecules may be selected that comprise at least 10, 15, 20, 25, 30, 35, 40, 50, 100, 150, 200, 250, 300 or more consecutive nucleotides of any of these or other portions of a bitter taste receptor nucleic acid molecule or a specific allele thereof, such as those disclosed herein. Thus, representative nucleic acid molecules might comprise at least 10 consecutive nucleotides of the bitter taste receptor nucleic acid coding sequence shown in any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 39, 41, or 45. More particularly, probes and primers in some embodiments are selected so that they overlap or reside adjacent to at least one of the indicated SNPs indicated in the Sequence Listing or in Figure 1 (Table of Bitter Taste Receptor Variants).

Purified: The term purified does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified nucleic acid preparation is one in which the specified protein is more enriched than the nucleic acid is in its generative environment, for instance within a cell or in a biochemical reaction

chamber. A preparation of substantially pure nucleic acid may be purified such that the desired nucleic acid represents at least 50% of the total nucleic acid content of the preparation. In certain embodiments, a substantially pure nucleic acid will represent at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, or at least 95% or more of the total nucleic acid content of the preparation.

Recombinant: A recombinant nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques.

RNA: A typically linear polymer of ribonucleic acid monomers, linked by phosphodiester bonds. Naturally occurring RNA molecules fall into three classes, messenger (mRNA, which encodes proteins), ribosomal (rRNA, components of ribosomes), and transfer (tRNA, molecules responsible for transferring amino acid monomers to the ribosome during protein synthesis). Total RNA refers to a heterogeneous mixture of all three types of RNA molecules.

Sequence identity: The similarity between two nucleic acid sequences, or two amino acid sequences, is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Homologs or orthologs of nucleic acid or amino acid sequences will possess a relatively high degree of sequence identity when aligned using standard methods. This homology will be more significant when the orthologous proteins or nucleic acids are derived from species which are more closely related (*e.g.*, human and chimpanzee sequences), compared to species more distantly related (*e.g.*, human and *C. elegans* sequences). Typically, orthologs are at least 50% identical at the nucleotide level and at least 50% identical at the amino acid level when comparing human orthologous sequences.

Methods of alignment of sequences for comparison are well known. Various programs and alignment algorithms are described in: Smith & Waterman, *Adv.*

Appl. Math. 2:482, 1981; Needleman & Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins & Sharp, Gene, 73:237-44, 1988; Higgins & Sharp, CABIOS 5:151-3, 1989; Corpet *et al.*, *Nuc. Acids Res.* 16:10881-90, 1988; Huang *et al. Computer Appls. Biosci.* 8, 155-65, 1992; and Pearson *et al.*, *Meth. Mol. Bio.* 24:307-31, 1994. Altschul *et al., J. Mol. Biol.* 215:403-10, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al., J. Mol. Biol.* 215:403-10, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. Each of these sources also provides a description of how to determine sequence identity using this program.

Homologous sequences are typically characterized by possession of at least 60%, 70%, 75%, 80%, 90%, 95% or at least 98% sequence identity counted over the full length alignment with a sequence using the NCBI Blast 2.0, gapped blastp set to default parameters. Queries searched with the blastn program are filtered with DUST (Hancock and Armstrong, *Comput. Appl. Biosci.* 10:67-70, 1994). It will be appreciated that these sequence identity ranges are provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions, as described under "specific hybridization."

Single Nucleotide Polymorphism (SNP): A single base (nucleotide) difference in a DNA sequence among individuals in a population.

Specific binding agent: An agent that binds substantially only to a defined target. Thus a protein-specific binding agent binds substantially only the specified protein. By way of example, as used herein, the term "X-protein specific binding agent" includes anti-X protein antibodies (and functional fragments thereof) and other agents (such as soluble receptors) that bind substantially only to the X protein (where "X" is a specified protein, or in some embodiments a specified domain or form of a protein, such as a particular allelic form of a protein).

Anti-X protein antibodies may be produced using standard procedures described in a number of texts, including Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). The determination that a particular agent binds substantially only to the specified protein may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988)). Western blotting may be used to determine that a given protein binding agent, such as an anti-X protein monoclonal antibody, binds substantially only to the X protein.

Shorter fragments of antibodies can also serve as specific binding agents. For instance, Fabs, Fvs, and single-chain Fvs (SCFvs) that bind to a specified protein would be specific binding agents. These antibody fragments are defined as follows: (1) Fab, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain; (2) Fab', the fragment of an antibody molecule obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule; (3) (Fab')2, the fragment of the antibody obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; (4) F(ab')2, a dimer of two Fab' fragments held together by two disulfide bonds; (5) Fv, a genetically engineered fragment containing the variable region of the light chain and the variable region of the heavy chain expressed as two chains; and (6) single chain antibody ("SCA"), a genetically engineered molecule containing the variable region of the light chain, the variable

region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule. Methods of making these fragments are routine.

Specific hybridization: Specific hybridization refers to the binding, duplexing, or hybridizing of a molecule only or substantially only to a particular nucleotide sequence when that sequence is present in a complex mixture (*e.g.* total cellular DNA or RNA). Specific hybridization may also occur under conditions of varying stringency.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing DNA used. Generally, the temperature of hybridization and the ionic strength (especially the Na+ concentration) of the hybridization buffer will determine the stringency of hybridization. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed by Sambrook *et al.* (In: *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989 ch. 9 and 11). By way of illustration only, a hybridization experiment may be performed by hybridization of a DNA molecule to a target DNA molecule which has been electrophoresed in an agarose gel and transferred to a nitrocellulose membrane by Southern blotting (Southern, *J. Mol. Biol.* 98:503, 1975), a technique well known in the art and described in Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989).

Traditional hybridization with a target nucleic acid molecule labeled with [^{32}P]-dCTP is generally carried out in a solution of high ionic strength such as 6 x SSC at a temperature that is 20-25° C below the melting temperature, T_m, described below. For Southern hybridization experiments where the target DNA molecule on the Southern blot contains 10 ng of DNA or more, hybridization is typically carried out for 6-8 hours using 1-2 ng/ml radiolabeled probe (of specific activity equal to 10^9 CPM/μg or greater). Following hybridization, the nitrocellulose filter is washed to remove background hybridization. The washing conditions should be as stringent as possible to remove background hybridization but to retain a specific hybridization signal.

The term T_m represents the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Because the target sequences are generally present in excess, at T_m 50% of the probes are occupied at equilibrium. The T_m of such a hybrid molecule may be estimated from the following equation (Bolton and McCarthy, *Proc. Natl. Acad. Sci. USA* 48:1390, 1962):

$$T_m = 81.5° C - 16.6(\log_{10}[Na^+]) + 0.41(\% G+C) - 0.63(\% \text{ formamide}) - (600/l)$$

where l = the length of the hybrid in base pairs.

This equation is valid for concentrations of Na^+ in the range of 0.01 M to 0.4 M, and it is less accurate for calculations of Tm in solutions of higher $[Na^+]$. The equation is also primarily valid for DNAs whose G+C content is in the range of 30% to 75%, and it applies to hybrids greater than 100 nucleotides in length (the behavior of oligonucleotide probes is described in detail in Ch. 11 of Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989).

Thus, by way of example, for a 150 base pair DNA probe derived from a cDNA (with a hypothetical % GC of 45%), a calculation of hybridization conditions required to give particular stringencies may be made as follows: For this example, it is assumed that the filter will be washed in 0.3 x SSC solution following hybridization, thereby: [Na+] = 0.045 M; %GC = 45%; Formamide concentration = 0; l = 150 base pairs; $Tm=81.5 - 16.6(\log_{10}[Na+]) + (0.41 \times 45) - (600/150)$; and so Tm = 74.4° C.

The T_m of double-stranded DNA decreases by 1-1.5° C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol.* 81:123, 1973). Therefore, for this given example, washing the filter in 0.3 x SSC at 59.4-64.4° C will produce a stringency of hybridization equivalent to 90%; that is, DNA molecules with more than 10% sequence variation relative to the target cDNA will not hybridize. Alternatively, washing the hybridized filter in 0.3 x SSC at a temperature of 65.4-68.4° C will yield a hybridization stringency of 94%; that is, DNA molecules with

more than 6% sequence variation relative to the target cDNA molecule will not hybridize. The above example is given entirely by way of theoretical illustration. It will be appreciated that other hybridization techniques may be utilized and that variations in experimental conditions will necessitate alternative calculations for stringency.

Stringent conditions may be defined as those under which DNA molecules with more than 25%, 15%, 10%, 6% or 2% sequence variation (also termed "mismatch") will not hybridize. Stringent conditions are sequence dependent and are different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5° C lower than the thermal melting point T_m for the specific sequence at a defined ionic strength and pH. An example of stringent conditions is a salt concentration of at least about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and a temperature of at least about 30° C for short probes (*e.g.* 10 to 50 nucleotides). Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide. For example, conditions of 5 X SSPE (750 mM NaCl, 50 mM Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30° C are suitable for allele-specific probe hybridizations.

A perfectly matched probe has a sequence perfectly complementary to a particular target sequence. The test probe is typically perfectly complementary to a portion (subsequence) of the target sequence. The term "mismatch probe" refers to probes whose sequence is deliberately selected not to be perfectly complementary to a particular target sequence.

Transcription levels can be quantitated absolutely or relatively. Absolute quantitation can be accomplished by inclusion of known concentrations of one or more target nucleic acids (for example control nucleic acids or with a known amount the target nucleic acids themselves) and referencing the hybridization intensity of unknowns with the known target nucleic acids (for example by generation of a standard curve).

Subject: Living, multicellular vertebrate organisms, a category that includes both human and veterinary subjects for example, mammals, birds and primates.

Transformed: A transformed cell is a cell into which has been introduced a nucleic acid molecule by molecular biology techniques. As used herein, the term transformation encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with viral vectors, transformation with plasmid vectors, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

Vector: A nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in a host cell, such as an origin of replication. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Hence "comprising A or B" means including A, or B, or A and B. It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

The following examples are provided to illustrate certain particular features and/or embodiments. These examples should not be construed to limit the invention to the particular features or embodiments described. In particular, other methods known to those of ordinary skill in the art can be substituted for specific methods

described herein. By way of example, additional methods for studying bitter taste receptors and compounds that interact therewith are described in PCT/US02/23172 (published as WO 03/008627), herein incorporated by reference in its entirety.

EXAMPLES

Example 1: Characterization of SNPs in the T2R (TAS2R) bitter taste receptor gene PTC

The ability to taste the substance phenylthiocarbamide (PTC) has been widely used for genetic and anthropological studies, but genetic studies have produced conflicting results and demonstrated complex inheritance for this trait. We have identified a small region on chromosome 7q that shows strong linkage disequilibrium between SNP markers and PTC taste sensitivity in unrelated subjects. This region contains a single gene that encodes a member of the TAS2R bitter taste receptor family. We identified three coding SNP's giving rise to five haplotypes in this gene worldwide. These haplotypes completely explain the bimodal distribution of PTC taste sensitivity, thus accounting for the inheritance of the classically defined taste insensitivity, and 55-85% of the variance in PTC sensitivity. Distinct phenotypes were associated with specific haplotypes, demonstrating the direct influence of this gene on PTC taste sensitivity, and that variant sites interact with each other within the encoded gene product.

Methods and Materials:

PTC phenotype determinations. Subjects began tasting a solution of 1 micromolar PTC (solution #14) and proceeded in 2-fold increasing concentration increments (solutions 13, 12, 11...) until a bitter taste was perceived. Subjects then performed a blinded sorting test containing 3 cups of PTC solution and 3 cups of water. Raw taste threshold was the most dilute solution at which the subject could correctly sort all 6 cups. We also included a quinine threshold measurement according to Blakeslee & Salmon (*Proc. Natl. Acad. Sci. USA* 21, 84,1935) to identify and exclude individuals with general deficits in bitter taste (aguesia). For dichotomous assignment of phenotype, we considered individuals unable to taste

PTC before solution #6, *i.e.* at concentrations less than 267 micromolar PTC, to be non-tasters. Although the classic method includes corrections for age and sex, analysis of our raw PTC taste threshold data indicated only a modest sex effect, with females more sensitive than males ($p = 0.00324$, proportion of variance explained = 5.1%). No effect of age on PTC scores was observed. As a result, raw PTC threshold scores were used for all analyses.

Research subjects. The Utah C.E.P.H. families were enrolled in conjunction with the Utah Genetic Reference Project under University of Utah IRB approved protocol #6090-96, and consisted of individuals of Northern European ancestry. Subjects in the NIH replication sample were enrolled under NIH/NINDS IRB approved protocol # DC-01-230, and were of European, Asian, African American, and Native American ancestry. Human Diversity Panel DNAs (sub-Saharan African, Asian, and Southwest Native American) and primate DNAs were obtained from the Coriell Cell Repository, Camden, NJ. The Utah sample consists of 27 families comprising 269 individuals; both haplotype and phenotype information was available for 180 of these individuals. The NIH replication sample of consisted of 85 unrelated individuals of known haplotype and phenotype; 51 were European, 5 Pakistani, 23 East Asian, and 6 African-American. One African-American is not considered in the analysis due to a rare AAV/AAI diplotype. His raw PTC score is 7.

Bioinformatics analyses. Bioinformatics analysis was performed with the NCBI Human Genome databases (http://www.ncbi.nlm.nih.gov/genome/guide/human) and the Celera Discovery System (http://cds.celera.com/cds). Gene finding was performed with BLASTX (http://www.ncbi.nlm.nih.gov/BLAST) and GENESCAN and FGENES software (GeneMachine, DIR, NIH, http://genome.nhgri.nih.gov/genemachine/). SNPs were developed using the SNP database (http://www.ncbi.nlm.nih.gov/SNP/)

PTC gene haplotyping. Haplotypes within the PTC gene were determined by performing genomic PCR to obtain a 1195 bp product containing all 3 variant sites, using primers as follows: F = 5' GCTTTGTGAGGAATCAGAGTTGT 3', R = 5' GAACGTACATTTACCTTTCTGCACT 3'. The mass PCR product from each

individual was cloned into TopoTA vector (Clonetech), and single colonies which contained a single amplified haplotype were picked and sequenced.

QTL linkage analysis. Quantitative trait linkage analysis was performed using SOLAR, Almassy and Blangero, *Am. J. Hum. Genet.* 62, 1198, 1998. The effect of PTC haplotypes on the linkage results was determined by performing two multipoint linkage analyses: one using the raw PTC scores and another using adjusted PTC scores, both with sex as a covariate. The first analysis excluded diplotypes as covariates, the second included them. For the latter, adjusted scores were obtained by subtracting off the mean of each diplotype group from the scores of individuals with that particular diplotype.

Haplotype effect analysis. The effect of the PTC haplotypes, as well as the covariates sex and age, on raw PTC scores was estimated simultaneously in a multivariate analysis using the program SOLAR[26]. SOLAR estimates the proportion of variance explained by a covariate (*e.g.*, the PTC diplotype) in the presence of background polygenic variance, in this case estimated from residual familial correlation in the phenotype. The program also takes into account non-independence of sib genotypes. The confirmation sample of unrelated individuals was analyzed using multiple linear regression with sex and age as covariates as well as Analysis of Variance.

GenBank. Human candidate taste receptor gene TAS2R38 (GenBank accession number AF494231) is identical to the sequence of the non-taster AVI form of the PTC gene, with the exception of nucleotide 557, which is an A (encoding Asn^{186}) in TAS2R38 and a T (encoding Ile^{186}) in PTC.

Material in this example was published as Kim *et al.*, Science 299:1221-1225, February 21, 2003, which publication is incorporated herein by reference in its entirety, including the supplemental material published on-line at www.scienemag.org/cgi/content/full/299/5610/121/DC1.

Results and Discussion

The inability to taste PTC (*Science* 73:4, 1931; Guo and Reed, *Ann. Hum. Biol.* 28:111, 2001) was long believed to be a simple Mendelian recessive trait (Snyder, *Science* 74:151, 1931; Levit and Soboleva, *J. Genetics* 30:389, 1935;

Blakeslee, *Proc. Acad. Natl. Acad. Sci. USA* 18:120, 1932; Lee, *Ohio J. Science* 34:337, 1934; Harris and Kalmus, *Ann. Eugenics, London* 15:24, 1949). Over time however, many reports emerged which contradicted this model (Falconer, *Ann. Eugenics* 13:211, 1946-47; Reddy and Rao, *Genet. Epidemiol.* 6:413, 1989; Olson *et al.*, *Genet. Epidemiol.* 6:423, 1989). Linkage studies have been equally conflicting. Initial studies provided very strong support for linkage to the KEL blood group antigen (later determined to reside on chromosome 7q3) (Chautard-Freire-Maia *et al.*, *Ann. Hum. Genet.* 38:191, 1974; Conneally *et al.*, *Hum. Hered.* 26:267, 1976), but other studies failed to provide significant support for this linkage (Spence *et al.*, *Hum. Genet.* 67:183, 1984). The only genome-wide linkage survey was performed with the related compound propyl-thiouracil. This study produced evidence for linkage to loci on chromosome 5p, and a suggestion of linkage to markers on chromosome 7q31, at a distance of ~35 cM from KEL (Reed *et al.*, *Am. J. Hum. Genet.* 64:1478, 1999).

We performed a genome-wide linkage analysis with the Utah C.E.P.H. families (Dausset *et al.*, *Genomics* 6:575, 1990; NIH/CEPH Collaborative Mapping Group, *Science* 258:67, 1992; Materials and methods are available as supporting material on *Science* Online.) using a blind sorting test to measure individual's PTC sensitivity thresholds (Materials and methods are available as supporting material on *Science* Online; Harris and Kalmus, *Ann. Eugenics, London* 15:24, 1949; Kalmus, *Ann. Hum. Genet.* 22:222, 1958), and demonstrated strong support for a major locus on chromosome 7q, close to KEL (Prodi *et al.*, *Am. J. Hum. Genet.* Suppl. 71(4):464, 2002; Drayna *et al.*, *Hum. Genet.* 112:567, 2003) with a critical region spanning approximately 4 Mb in the region of D7S661, with a maximum lod score of 8.85 (Drayna *et al.*, *Hum. Genet.* 112:567, 2003).

Bioinformatic analyses (Materials and methods are available as supporting material on *Science* Online.) indicated the ~4 Mb region on chromosome 7q contains over 150 genes, including the KEL blood group antigen, confirming previous linkage studies (Chautard-Freire-Maia *et al.*, *Ann. Hum. Genet.* 38:191, 1974; 13. Conneally *et al.*, *Hum. Hered.* 26:267, 1976). In addition, this region contains a number of TAS2R bitter taste receptor genes (Adler *et al.*, *Cell* 100:693, 2000) and odorant receptor-like genes (Buck and Axel, *Cell* 65:175, 1991). All

TAS2R's (9 genes) and OR-like genes (7 genes) were evaluated as candidates by sequencing the entire single coding exon, the 3' UTR, and 300 bp upstream in individuals within families showing linkage to chromosome 7q, and numerous sequence variants were observed (Ewing *et al.*, *Genome Res.* 8:175, 1998; 26. Gordon *et al.*, *Genome Res.* 8:195, 1998. Seqman (DNA STAR, Madison, WI)). One of these variants demonstrated strong association with taste phenotype across different C.E.P.H. families (chi-square $p<10^{-10}$), suggesting it may be the functional change or close to the functional change(s). To more fully understand linkage and LD relationships in this region, we performed further analysis by means of 50 SNPs at an average spacing of 50 kb across this interval. These SNPs revealed crossover breakpoints in the Utah C.E.P.H. families that reduced the minimal region to 2.6 Mb.

Using these 50 SNP's, strong LD was observed between taster status and markers in only one portion of this 2.6 Mb interval. This was observed initially in the chromosome 7-linked families (12 families containing 107 individuals) and subsequently in unrelated non-tasters from both the C.E.P.H. sample (an additional 8 individuals) and in a second replication population (the NIH sample, 15 non-taster and 14 taster Caucasians, 7 non-taster and 9 taster East Asians). Significant LD was observed across a 150 kb region, extending from approximately 139,835,000 to 139,981,000 bp on the chromosome 7 genomic sequence (http://www.ncbi.nlm.nih.gov/genome/guide/human). In the NIH sample of 45 individuals, analysis of chi-square (equivalent to r^2) and delta statistics showed clear peak values for each measure within the BAC RP11-707F14 (AC073647.9) ($p < 10^{-10}$), at the identical location in the Caucasian and East Asian subgroups as well as for the Mantel-Haenszel combined chi-square. In a group of 37 unrelated non-taster individuals (12 Utah individuals and 25 individuals from the NIH sample who collectively had the poorest PTC sensitivities), the physical distance over which these individuals carried unambiguous haplotypes sharing the same SNP alleles extended an average of 61 kb, with the minimal shared region extending from 42,445 bp to 72,141 bp in this BAC, a distance of 29,696 bp. Bioinformatic and gene prediction analyses revealed that the only gene in this 29.7 kb interval was the TAS2R bitter receptor gene in which we originally identified strong LD.

This gene, which we have designated PTC, consists of 1002 bp in a single exon, encoding a 7 transmembrane domain, G-protein-coupled receptor that shows 30% amino acid identity with human TAS2R7, the most closely related member of this family. Within this gene, we identified 3 common SNPs, all of which result in amino acid changes in the protein (Table 1). The A49P variant demonstrated a strong association overall with taster status in the Utah sample (Table 2), and an even stronger association in the NIH replication sample (Table 2). The association of taster status with the val262 allele was similarly strong in both the Utah and NIH samples (Table 2). To better understand the effect of these SNP's, we investigated the haplotypes in this gene.

Haplotype analysis in the Utah and NIH samples revealed two predominant haplotypes at the three SNPs in this gene. Named in the order of the three SNPs (A49P, V262A, and I296V), the non-taster haplotype AVI and taster haplotype PAV accounted for 47% and 49% of all haplotypes respectively in the European sample, and 30% and 70% respectively in the East Asian sample. Europeans also possessed the presumed recombinant taster haplotype AAV at a frequency of 3%. The haplotype association with taster status was more definitive than for individual SNP's; the strongest association with non-taster status is for the AVI homozygote, followed by the compound heterozygote AVI/AAV (Table 3).

Due to the broad and continuous distribution of PTC sensitivity in the population, we went on to analyze PTC scores as a quantitative trait. There was a consistent and significant difference in PTC scores between diplotypes in both the Utah and the NIH samples, consistent across racial groups. PAV homozygotes had the highest mean PTC scores (Utah: 10.69, NIH: 10.00), PAV heterozygotes had slightly but significantly lower mean PTC scores (Utah: 9.65, NIH: 8.81) than the PAV homozygotes (Utah sample: $\chi^2 = 8.41$, $p = 0.0037$, NIH replication sample: $t = 3.29$, $p = 0.0017$). AVI homozygotes had the lowest mean PTC scores (Utah: 4.31, NIH: 1.86). Thus the taster PAV form of the gene displays a heterozygote effect, with two copies conferring greater PTC sensitivity than a single copy. The difference in mean PTC score between the rare AAV/AVI heterozygotes and the AVI homozygotes was significant in the NIH sample ($t = 5.44$, $p = 5.41 \times 10^{-5}$) and tended toward significance in the Utah family sample ($\chi^2 = 2.39$, $p = 0.122$).

PAV/AAV heterozygotes were not significantly different from PAV/AVI heterozygotes ($\chi^2 = 0.58$, $p = 0.45$).

Differences in PTC score by diplotype in the Utah families were also highly significant in a multivariate analysis ($\chi^2 = 148.95$, $p < 10^{-33}$) (18). Sex and the haplotype effect explain 59.7% of the total variance in PTC scores. Analysis of variance of the NIH sample confirmed these results ($F = 152.73$, $p < 10^{-32}$), with 84.8% of the variance explained by the haplotype effect. The differences were also significant in both the Caucasian subgroup of the replication sample ($F = 78.60$, $p < 10^{-18}$) and the East Asian subgroup ($F = 139.02$, $p < 10^{-11}$).

The bimodal distribution of PTC scores is a combination of the underlying distributions of the PTC diplotypes, i.e. genotypes at multiple variable sites with consideration of haplotype. The appearance of bimodality is driven by the distribution of the common AVI homozygote, PAV/AVI heterozygote and PAV homozygote diplotypes. The mode of inheritance of PTC taste sensitivity has been a subject of controversy (Guo and Reed, *Ann. Hum. Biol.* 28:111, 2001; Reddy and Rao, *Genet. Epidemiol.* 6:413, 1989; Olson *et al.*, *Genet. Epidemiol.* 6:423, 1989). To determine whether there was evidence for additional genetic contributions to PTC score, we examined the heritability in subsets of the Utah sample. In the subgroups which were large enough to give accurate estimates, heritability was 0.26 ± 0.19 (83 subjects in 20 families) in the PAV/AVI subgroup, and 0.50 ± 0.33 in the AVI/AVI subgroup (46 subjects in 17 families). The increase in heritability in the loss of function diplotype group (AVI/AVI) indicates that there may be other genetic factors that interact with PTC and can restore some measure of taste sensitivity in this group. For Caucasians and East Asians, our results are largely consistent with a model of a major recessive QTL modified either by a polygenic (Reddy and Rao, *Genet. Epidemiol.* 6:413, 1989) or single locus (Olson *et al.*, *Genet. Epidemiol.* 6:423, 1989) residual background effect.

Due to the high frequency of the PAV and AVI haplotypes in the population, we sought to determine which haplotype represents the original form of the PTC gene. We sequenced this gene in 6 primate species: humans and one individual each from chimpanzee, lowland gorilla, orangutan, crab-eating macaque (an old world monkey), and black-handed spider monkey (a new world monkey), representing

over 25 million years of evolutionary divergence. All of the non-human primates were homozygous for the PAV form, indicating that the AVI form arose in humans after the time they diverged from the nearest common primate ancestors.

Five different haplotypes were observed worldwide (Table 4). In Europeans and Asians, the taster haplotype PAV and the non-taster haplotype AVI make up the vast majority of haplotypes present. Two additional haplotypes, PVI and AAI, were observed only in individuals of sub-Saharan African ancestry, consistent with other reports of increased gene haplotype diversity in this population (Stephens *et al.*, *Science* 293:489, 2001). The common non-taster AVI haplotype was observed in all populations except Southwest Native Americans, who were exclusively homozygous for the PAV haplotype, consistent with the reported low frequency of non-tasters in this population (Guo and Reed, *Ann. Hum. Biol.* 28:111, 2001). Thus overall, the worldwide distribution of these haplotypes is consistent with the large anthropologic literature on the distribution of this phenotype (Boyd, "Genetics & the Races of Man. An introduction to modern physical anthropology." Little Brown and Company, Boston, 1950; Tills *et al.*, "The Distribution of Human Blood Groups and other Polymorphisms," Supplement, 1st Edition. Oxford University Press, Oxford, 1983).

The amino acid substitutions in the PTC protein may affect the function of this protein in several ways. Position 49 resides in the predicted first intracellular loop, and this SNP represents a major amino acid alteration, proline in tasters to alanine in non-tasters. The SNP's at positions 262, in the predicted 6th transmembrane domain, and position 296, in the predicted 7th transmembrane domain specify relatively conserved amino acid changes, alanine to valine and valine to isoleucine, respectively. Based on phenotype data, we hypothesize that the substitutions at positions 49 and 262 significantly alter the biochemical function of this protein, while the substitution at position 296 modifies the function more subtly. These alterations could affect coupling to its cognate G proteins on the intracellular side of the plasma membrane, as has been observed for other variants in the first intracellular loop (Nabhan*et al.*, *Biochem. Biophys. Res. Comm.* 212:1015, 1995; O'Dowd *et al.*, *J. Biol. Chem.* 263:15985, 1988), or in other portions of these proteins (G protein receptor database: www.gpcr.org, www.grap.fagmed.uit.no).

Given that PTC and other compounds which contain the N-C=S moiety are both bitter and toxic in large doses, it will be of interest to determine how the non-taster allele rose to such high frequency, especially in the European population.

Substantial variation in taste sensitivity exists in humans (Blakeslee and Salmon, *Proc. Natl. Acad. Sci. USA* 21:84, 1935), and given the great degree of sequence diversity and variation in bitter taste receptor genes (Ueda *et al.*, *Biochem. Biophys. Res. Comm.* 285:147, 2001), we hypothesize much of this phenotypic variation is genetic in origin. Understanding the nature of this variation, especially variation in bitter taste, and its relationship to diet and other behaviors such as smoking may have important implications for human health (Tepper, *Am. J. Hum. Genet.* 63:1271, 1998; Enoch *et al.*, *Addictive Behav.* 26:399, 2001).

Table 1 Polymorphisms within the PTC gene

Position (b.p.)	Position (a.a.)	Allele	Frequency	AA encoded
145	49	C	.48	Pro
		G	.51	Ala
785	262	C	.38	Ala
		T	.62	Val
886	296	G	.38	Val
		A	.62	Ile

Table 2. The effect of homozygosity for SNPs on phenotype

Homozygous SNP	Sample	No. of subjects (total no.)		χ^2	*P* value
		Non-tasters	Tasters		
Ala 49	Utah	48 (51)	21 (129)	27.23	1.81×10^{-7}
	NIH	22 (23)	3 (61)	72.74	1.61×10^{-16}
Val 262	Utah	38 (51)	14 (129)	23.40	1.10×10^{-6}
	NIH	21 (23)	0 (61)	74.44	6.83×10^{-17}

* The third SNP, I296V, was in complete linkage disequilibrium with V262A (and thus gave identical results to V262A) except in one African-American subject.

Table 3. Haplotype association with taste phenotypes

Haplotypes	Sample	No. of subjects	
		Non-tasters	Tasters
AVI / AVI	Utah	38	14
	NIH	21	0
AVI / AAV	Utah	10	7
	NIH	1	3
* / PAV	Utah	3	108
	NIH	1	58

* indicates any haplotype found in the sample. No AAV homozygotes were observed in either sample.

Table 4. Frequency of PTC gene haplotypes in populations worldwide

Haplotype	European (n = 200)	West Asian (n = 22)	East Asian (n = 54)	African (n = 24)	S.W. Native American (n = 18)
AVI	0.47	0.67	0.31	0.25	
AAV	0.03			0.04	
AAI				0.17	
PAV	0.49	0.33	0.69	0.50	1.00
PVI				0.04	

Example 2: Identification of SNPs in Other T2R Bitter Taste Receptors

Common allelic variants of a member of the TAS2R bitter taste receptor gene family underlie variation in the ability to taste phenylthiocarbamide (PTC). To extend these results to other bitter receptors, we have sequenced 22 of the 24 known TAS2R genes in a series of populations worldwide, including Hungarians, Japanese, Cameroonians, Pygmies, and South American Indians. This example provides description of this analysis, which was used to generate a comprehensive collection of single nucleotide polymorphisms in human T2R putative bitter taste receptors.

Using conventional methods, members of the human T2R family of putative bitter taste receptors were analyzed for the presence of SNPs. All SNPs were identified and analyzed by DNA sequencing. Genomic DNA encoding each receptor was PCR'd using standard methods, and the products cycle sequenced with dye terminators using a Big Dye terminator kit from ABI. Products of the sequencing reactions were analyzed on an ABI 3730x1 DNA Analyzer using the manufactures' recommendations. Other sequencing techniques would be equally applicable to detecting SNPs in these genes.

The results of the comprehensive sequencing are presented in Figure 1; specific individual variants are also described in the attached Sequence Listing. Figure 1 shows, in addition to those SNPs confirmed or identified by our sequencing reaction, all SNPs found in dbSNP for these genes.

All 22 TAS2R genes contain common SNP's within their coding sequence, and we identified an average of 4.4 SNP's per TAS2R gene. Fifteen variants listed in dbSNP were not observed to be polymorphic in our sample. However, many novel SNPs were identified; these are indicated with the "new" designation in Figure 1. Of the SNP's we observed, 77% cause an amino acid substitution in the encoded receptor protein, giving rise to a very high degree of receptor protein variation in the population. Four SNP's specify one allele that introduces an in-frame stop codon in the gene. Some of the SNP's were observed only in individuals of sub-Saharan African origin, and overall African samples displayed higher diversity of alleles. This is consistent with the view that the majority of human genetic variation resides within older African populations, and a fraction of this variation emerged and subsequently spread across the remainder of the world.

Example 3: Detecting Single Nucleotide Alterations

T2R bitter taste receptor single nucleotide alterations, whether categorized as SNPs or new mutations can be detected by a variety of techniques in addition to merely sequencing the target sequence. Constitutional single nucleotide alterations can arise either from new germline mutations, or can be inherited from a parent who possesses a SNP or mutation in their own germline DNA. The techniques used in evaluating either somatic or germline single nucleotide alterations include

hybridization using allele specific oligonucleotides (ASOs) (Wallace *et al.*, CSHL Symp. Quant. Biol. 51:257-261, 1986; Stoneking *et al.*, *Am. J. Hum. Genet.* 48:370-382, 1991), direct DNA sequencing (Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81:1991-1995, 1988), the use of restriction enzymes (Flavell *et al.*, *Cell* 15:25, 1978; Geever *et al.*, 1981), discrimination on the basis of electrophoretic mobility in gels with denaturing reagent (Myers and Maniatis, Cold Spring Harbor Symp. Quant. Biol. 51:275-284, 1986), RNase protection (Myers *et al.*, *Science* 230:1242, 1985), chemical cleavage (Cotton *et al.*, *Proc. Natl. Acad. Sci. USA* 85:4397-4401, 1985), and the ligase-mediated detection procedure (Landegren *et al.*, *Science* 241:1077, 1988).

Allele-specific oligonucleotide hybridization (ASOH) involves hybridization of probes to the sequence, stringent washing, and signal detection. Other new methods include techniques that incorporate more robust scoring of hybridization. Examples of these procedures include the ligation chain reaction (ASOH plus selective ligation and amplification), as disclosed in Wu and Wallace (*Genomics* 4:560-569, 1989); mini-sequencing (ASOH plus a single base extension) as discussed in Syvanen (*Meth. Mol. Biol.* 98:291-298, 1998); and the use of DNA chips (miniaturized ASOH with multiple oligonucleotide arrays) as disclosed in Lipshutz *et al.* (*BioTechniques* 19:442-447, 1995). Alternatively, ASOH with single- or dual-labeled probes can be merged with PCR, as in the 5'-exonuclease assay (Heid *et al.*, *Genome Res.* 6:986-994, 1996), or with molecular beacons (as in Tyagi and Kramer, *Nat. Biotechnol.* 14:303-308, 1996).

Another technique is dynamic allele-specific hybridization (DASH), which involves dynamic heating and coincident monitoring of DNA denaturation, as disclosed by Howell *et al.* (*Nat. Biotech.* 17:87-88, 1999). A target sequence is amplified by PCR in which one primer is biotinylated. The biotinylated product strand is bound to a streptavidin-coated microtiter plate well, and the non-biotinylated strand is rinsed away with alkali wash solution. An oligonucleotide probe, specific for one allele, is hybridized to the target at low temperature. This probe forms a duplex DNA region that interacts with a double strand-specific intercalating dye. When subsequently excited, the dye emits fluorescence proportional to the amount of double-stranded DNA (probe-target duplex) present.

The sample is then steadily heated while fluorescence is continually monitored. A rapid fall in fluorescence indicates the denaturing temperature of the probe-target duplex. Using this technique, a single-base mismatch between the probe and target results in a significant lowering of melting temperature (T_m) that can be readily detected.

Oligonucleotides specific to normal or allelic sequences can be chemically synthesized using commercially available machines. These oligonucleotides can then be labeled radioactively with isotopes (such as ^{32}P) or non-radioactively, with tags such as biotin (Ward and Langer *et al., Proc. Natl. Acad. Sci. USA* 78:6633-6657, 1981), and hybridized to individual DNA samples immobilized on membranes or other solid supports by dot-blot or transfer from gels after electrophoresis. These specific sequences are visualized by methods such as autoradiography or fluorometric (Landegren *et al., Science* 242:229-237, 1989) or colorimetric reactions (Gebeyehu *et al., Nucleic Acids Res.* 15:4513-4534, 1987). Using an ASO specific for a normal allele, the absence of hybridization would indicate a mutation in the particular region of the gene, or a deleted gene. In contrast, if an ASO specific for a mutant allele hybridizes to a sample then that would indicate the presence of a mutation in the region defined by the ASO.

A variety of other techniques can be used to detect the mutations or other variations in DNA. Merely by way of example, see U.S. Patents No. 4,666,828; 4,801,531; 5,110,920; 5,268,267; 5,387,506; 5,691,153; 5,698,339; 5,736,330; 5,834,200; 5,922,542; and 5,998,137 for such methods. Additional methods include fluorescence polarization methods such as those developed by Pui Kwok and colleagues (see, *e.g.*, Kwok, Hum. Mutat., 19(4):315-23, 2002), microbead methods such as those developed by Mark Chee at Illumina (see, *e.g.*, Oliphant *et al., Biotechniques*. 2002 Jun;Suppl:56-8, 60-61, Shen *et al., Genet. Eng. News*, 23(6), 2003), and mass spectrophotometery methods such as those being developed at Sequenom (www.sequenom.com) (see, *e.g.*, Jurinke *et al., Methods Mol Biol.* 187:179-92, 2002; Amexis *et al., Proc Natl Acad Sci USA* 98(21):12097-102, 2001; Jurinke *et al., Adv Biochem Eng Biotechnol.* 2002;77:57-74; Storm *et al., Meth. Mol. Biol.*, 212:241 262, 2002; Rodi *et al., BioTechniques.*, 32:S62 S69, 2002); USPN 6,300,076; and WO9820166).

Example 4: Differentiation of Individuals Homozygous versus Heterozygous for Activating Mutation(s)

Since it is believed that the haplotype of any taste receptor can influence the perception of taste by a subject, it may sometimes be beneficial to determine whether a subject is homozygous or heterozygous for SNPs within any one or more of the T2R bitter taste receptors described herein.

By way of example, the oligonucleotide ligation assay (OLA), as described at Nickerson *et al.* (*Proc. Natl. Acad. Sci. USA* 87:8923-8927, 1990), allows the differentiation between individuals who are homozygous versus heterozygous for alleles indicated in Figure 1. This feature allows one to rapidly and easily determine whether an individual is homozygous for at least one tyrosine kinase activating mutation, which condition is linked to a relatively high predisposition to developing neoplastic disease and/or an increased likelihood of having a tumor. Alternatively, OLA can be used to determine whether a subject is homozygous for either of these mutations.

As an example of the OLA assay, when carried out in microtiter plates, one well is used for the determination of the presence of the T2R bitter taste allele in the T2R1 gene that contains an A at nucleotide position 332 (numbering from SEQ ID NO: 1) and a second well is used for the determination of the presence of the T2R bitter taste allele in the same gene that contains a G at that nucleotide position in the alternate allele sequence. Thus, the results for an individual who is heterozygous for the mutation will show a signal in each of the A and G wells.

Example 4: Bitter Taste Profiles

With the provision herein of specific SNPs within the family of bitter taste receptors that are linked to bitter taste sensitivity to one or more bitter compounds, genetic profiles that provide information on the bitter taste perception of a subject are now enabled.

Bitter taste-related genetic profiles comprise the distinct and identifiable pattern of alleles or haplotypes, or sets of of alleles or haplotypes, of the SNPs in bitter taste receptor molecules identified herein. The set of bitter taste receptors

analyzed in a particular profile will usually include at least one of the following: T2R1, T2R3, T2R4, T2R5, T2R7, T2R8, T2R9, T2R10, T2R13, T2, R14, T2R16, TwR38, TwR39, T2R40, T2R40, T2R43, T2R44, T2R45, T2R46, T2R47, T2R48, T2R49, T2R50, or T2R60.

By way of example, any subset of the molecules listed in Figure 1 (or corresponding to the molecules in this list) may be included in a single bitter taste profile. Specific examples of such subsets include those molecules that show a SNP that introduces a stop codon (*e.g.*, the variant of T2R44 at position 843; the variant of T2R46 at position 749 or 86, or the variant of T2R48 at position 885); that show a novel SNP (*e.g.*, those T2R genes with a "new" SNP indicated in Figure 1); and so forth. Alternatively, gene profiles may be further broken down by the type of bitter taste receptors included in the profile, for instance, those which all occur on a single chromosome (*e.g.*, CH 5, 7, or 12).

The alleles/haplotypes of each bitter taste receptor included in a specific profile can be determined in any of various known ways, including specific methods provided herein. One particular contemplated method for detecting and determining the genotype and/or haplotype of multiple bitter taste receptors employs an array of allele-specific oligonucleotides which are used for qualitative and/or quantitative hybridization detection of the presence of specific alleles or SNPs in a sample from a subject.

Optionally, a subject's bitter taste profile can be correlated with one or more appropriate inhibitors or blockers of bitter taste, or other compounds that influence the ability of a subject to perceive a taste, which may be correlated with a control (or set of control) profile(s) condition linked to or associated with, for instance, sensitivity to one or a set of bitter compounds. Optionally, the subject's bitter taste profile can be correlated with one or more appropriate treatments, for instance, treatments with compounds that inhibit or enhance the activity of one or more of the bitter taste alleles identified in the profile, or compositions in which the bitter taste of a component is specifically masked by a blocker that is added based on the information in the profile.

Example 6: Expression of T2R Bitter Taste Receptor Variant Polypeptides

The expression and purification of proteins, such as a T2R bitter taste receptor variant protein, can be performed using standard laboratory techniques, though these techniques are preferentially adapted to be fitted to express the T2R proteins. Examples of such method adaptations are discussed or referenced herein. After expression, purified protein may be used for functional analyses, antibody production, diagnostics, and patient therapy. Furthermore, the DNA sequences of the T2R bitter taste receptor variant cDNAs can be manipulated in studies to understand the expression of the gene and the function of its product. Variant or allelic forms of a human T2R bitter taste receptor genes may be isolated based upon information contained herein, and may be studied in order to detect alteration in expression patterns in terms of relative quantities, tissue specificity and functional properties of the encoded T2R bitter taste receptor variant protein (*e.g.*, influence on perception of taste). Partial or full-length cDNA sequences, which encode for the subject protein, may be ligated into bacterial expression vectors. Methods for expressing large amounts of protein from a cloned gene introduced into *Escherichia coli* (*E. coli*) or more preferably baculovirus/Sf9 cells may be utilized for the purification, localization and functional analysis of proteins. For example, fusion proteins consisting of amino terminal peptides encoded by a portion of a gene native to the cell in which the protein is expressed (*e.g.*, a *E. coli lacZ* or *trpE* gene for bacterial expression) linked to a T2R bitter taste receptor variant protein may be used to prepare polyclonal and monoclonal antibodies against these proteins. Thereafter, these antibodies may be used to purify proteins by immunoaffinity chromatography, in diagnostic assays to quantitate the levels of protein and to localize proteins in tissues and individual cells by immunofluorescence.

Intact native protein may also be produced in large amounts for functional studies. Methods and plasmid vectors for producing fusion proteins and intact native proteins in culture are well known in the art, and specific methods are described in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 17, CSHL, New York, 1989). Such fusion proteins may be made in large amounts, are easy to purify, and can be used to elicit antibody response. Native proteins can be produced in bacteria by placing a strong, regulated promoter and an efficient

ribosome-binding site upstream of the cloned gene. If low levels of protein are produced, additional steps may be taken to increase protein production; if high levels of protein are produced, purification is relatively easy. Suitable methods are presented in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989) and are well known in the art. Often, proteins expressed at high levels are found in insoluble inclusion bodies. Methods for extracting proteins from these aggregates are described by Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 17, CSHL, New York, 1989). Vector systems suitable for the expression of *lacZ* fusion genes include the pUR series of vectors (Ruther and Muller-Hill, *EMBO J.* 2:1791, 1983), pEX1-3 (Stanley and Luzio, *EMBO J.* 3:1429, 1984) and pMR100 (Gray *et al.*, *Proc. Natl. Acad. Sci. USA* 79:6598, 1982). Vectors suitable for the production of intact native proteins include pKC30 (Shimatake and Rosenberg, *Nature* 292:128, 1981), pKK177-3 (Amann and Brosius, *Gene* 40:183, 1985) and pET-3 (Studiar and Moffatt, *J. Mol. Biol.* 189:113, 1986).

Fusion proteins may be isolated from protein gels, lyophilized, ground into a powder and used as an antigen. The DNA sequence can also be transferred from its existing context to other cloning vehicles, such as other plasmids, bacteriophages, cosmids, animal viruses and yeast artificial chromosomes (YACs) (Burke *et al.*, *Science* 236:806-812, 1987). These vectors may then be introduced into a variety of hosts including somatic cells, and simple or complex organisms, such as bacteria, fungi (Timberlake and Marshall, *Science* 244:1313-1317, 1989), invertebrates, plants (Gasser and Fraley, *Science* 244:1293, 1989), and animals (Pursel *et al.*, *Science* 244:1281-1288, 1989), which cell or organisms are rendered transgenic by the introduction of the heterologous cDNA.

For expression in mammalian cells, the cDNA sequence may be ligated to heterologous promoters, such as the simian virus (SV) 40 promoter in the pSV2 vector (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981), and introduced into cells, such as monkey COS-1 cells (Gluzman, *Cell* 23:175-182, 1981), to achieve transient or long-term expression. The stable integration of the chimeric gene construct may be maintained in mammalian cells by biochemical selection, such as neomycin (Southern and Berg, *J. Mol. Appl. Genet.* 1:327-341,

1982) and mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981).

DNA sequences can be manipulated with standard procedures such as restriction enzyme digestion, fill-in with DNA polymerase, deletion by exonuclease, extension by terminal deoxynucleotide transferase, ligation of synthetic or cloned DNA sequences, site-directed sequence-alteration via single-stranded bacteriophage intermediate or with the use of specific oligonucleotides in combination with PCR or other *in vitro* amplification.

The cDNA sequence (or portions derived from it) or a mini gene (a cDNA with an intron and its own promoter) may be introduced into eukaryotic expression vectors by conventional techniques. These vectors are designed to permit the transcription of the cDNA in eukaryotic cells by providing regulatory sequences that initiate and enhance the transcription of the cDNA and ensure its proper splicing and polyadenylation. Vectors containing the promoter and enhancer regions of the SV40 or long terminal repeat (LTR) of the Rous Sarcoma virus and polyadenylation and splicing signal from SV40 are readily available (Mulligan *et al., Proc. Natl. Acad. Sci. USA* 78:1078-2076, 1981; Gorman *et al., Proc. Natl. Acad. Sci USA* 78:6777-6781, 1982). The level of expression of the cDNA can be manipulated with this type of vector, either by using promoters that have different activities (for example, the baculovirus pAC373 can express cDNAs at high levels in *S. frugiperda* cells (Summers and Smith, In *Genetically Altered Viruses and the Environment*, Fields *et al.* (Eds.) 22:319-328, CSHL Press, Cold Spring Harbor, New York, 1985) or by using vectors that contain promoters amenable to modulation, for example, the glucocorticoid-responsive promoter from the mouse mammary tumor virus (Lee *et al., Nature* 294:228, 1982). The expression of the cDNA can be monitored in the recipient cells 24 to 72 hours after introduction (transient expression).

In addition, some vectors contain selectable markers such as the *gpt* (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981) or *neo* (Southern and Berg, *J. Mol. Appl. Genet.* 1:327-341, 1982) bacterial genes. These selectable markers permit selection of transfected cells that exhibit stable, long-term expression of the vectors (and therefore the cDNA). The vectors can be maintained in the cells as episomal, freely replicating entities by using regulatory elements of

viruses such as papilloma (Sarver *et al., Mol. Cell Biol.* 1:486, 1981) or Epstein-Barr (Sugden *et al., Mol. Cell Biol.* 5:410, 1985). Alternatively, one can also produce cell lines that have integrated the vector into genomic DNA. Both of these types of cell lines produce the gene product on a continuous basis. One can also produce cell lines that have amplified the number of copies of the vector (and therefore of the cDNA as well) to create cell lines that can produce high levels of the gene product (Alt *et al., J. Biol. Chem.* 253:1357, 1978).

The transfer of DNA into eukaryotic, in particular human or other mammalian cells, is now a conventional technique. The vectors are introduced into the recipient cells as pure DNA (transfection) by, for example, precipitation with calcium phosphate (Graham and vander Eb, *Virology* 52:466, 1973) or strontium phosphate (Brash *et al., Mol. Cell Biol.* 7:2013, 1987), electroporation (Neumann *et al., EMBO J* 1:841, 1982), lipofection (Felgner *et al., Proc. Natl. Acad. Sci USA* 84:7413, 1987), DEAE dextran (McCuthan *et al., J. Natl. Cancer Inst.* 41:351, 1968), microinjection (Mueller *et al., Cell* 15:579, 1978), protoplast fusion (Schafner, *Proc. Natl. Acad. Sci. USA* 77:2163-2167, 1980), or pellet guns (Klein *et al., Nature* 327:70, 1987). Alternatively, the cDNA, or fragments thereof, can be introduced by infection with virus vectors. Systems are developed that use, for example, retroviruses (Bernstein *et al., Gen. Engr'g* 7:235, 1985), adenoviruses (Ahmad *et al., J. Virol.* 57:267, 1986), or Herpes virus (Spaete *et al., Cell* 30:295, 1982). Tyrosine kinase encoding sequences can also be delivered to target cells *in vitro* via non-infectious systems, for instance liposomes.

These eukaryotic expression systems can be used for studies of T2R bitter taste receptor variant encoding nucleic acids and mutant forms of these molecules, T2R bitter taste receptor variant proteins and mutant forms of these proteins. The eukaryotic expression systems may also be used to study the function of the normal complete protein, specific portions of the protein, or of naturally occurring or artificially produced mutant proteins.

Using the above techniques, the expression vectors containing a T2R gene sequence or cDNA, or fragments or variants or mutants thereof, can be introduced into human cells, mammalian cells from other species or non-mammalian cells as desired. The choice of cell is determined by the purpose of the treatment. For

example, monkey COS cells (Gluzman, *Cell* 23:175-182, 1981) that produce high levels of the SV40 T antigen and permit the replication of vectors containing the SV40 origin of replication may be used. Similarly, Chinese hamster ovary (CHO), mouse NIH 3T3 fibroblasts or human fibroblasts or lymphoblasts may be used.

The present disclosure thus encompasses recombinant vectors that comprise all or part of the T2R bitter taste receptor variant gene or cDNA sequences, for expression in a suitable host. The T2R bitter taste receptor DNA is operatively linked in the vector to an expression control sequence in the recombinant DNA molecule so that a T2R bitter taste receptor polypeptide can be expressed. The expression control sequence may be selected from the group consisting of sequences that control the expression of genes of prokaryotic or eukaryotic cells and their viruses and combinations thereof. The expression control sequence may be specifically selected from the group consisting of the *lac* system, the *trp* system, the *tac* system, the *trc* system, major operator and promoter regions of phage lambda, the control region of fd coat protein, the early and late promoters of SV40, promoters derived from polyoma, adenovirus, retrovirus, baculovirus and simian virus, the promoter for 3-phosphoglycerate kinase, the promoters of yeast acid phosphatase, the promoter of the yeast alpha-mating factors and combinations thereof.

One highly successful method of expressing T2R's to date is to engineer an amino-terminal portion of rhodopsin (*e.g.*, the first 26 amino acids thereof) onto the amino terminal end and express the resultant fusion protein, for instance in a baculovirus/Sf9 cell system. By way of example, methods for expressing T2Rs *in vitro* are described in Chandrashekar *et al.* (Cell 100:703-711, 2000), which is incorporated herein by reference in its entirety. See also Vince *et al.*, *PNAS* 99:2392-2397, 2002.

The host cell, which may be transfected with the vector of this disclosure, may be selected from the group consisting of *E. coli*, *Pseudomonas*, *Bacillus subtilis*, *Bacillus stearothermophilus* or other bacilli; other bacteria; yeast; fungi; insect; mouse or other animal; or plant hosts; or human tissue cells.

It is appreciated that for mutant or variant T2R bitter taste receptor DNA sequences, similar systems are employed to express and produce the mutant product.

In addition, fragments of a T2R bitter taste receptor protein can be expressed essentially as detailed above. Such fragments include individual T2R bitter taste receptor protein domains or sub-domains, as well as shorter fragments such as peptides. T2R bitter taste receptor protein fragments having therapeutic properties may be expressed in this manner also, including for instance substantially soluble fragments.

Example 7: Production of Protein Specific Binding Agents

Monoclonal or polyclonal antibodies may be produced to either a wildtype or reference T2R bitter taste receptor protein or specific allelic forms of these proteins, for instance particular portions that contain a differential amino acid encoded by a SNP and therefore may provide a distinguishing epitope. Optimally, antibodies raised against these proteins or peptides would specifically detect the protein or peptide with which the antibodies are generated. That is, an antibody generated to the specified bitter taste receptor protein or a fragment thereof would recognize and bind that protein and would not substantially recognize or bind to other proteins found in human cells. In some embodiments, an antibody is specific for (or measurably preferentially binds to) an epitope in a variant protein (*e.g.*, an allele of a T2R bitter taste receptor as described herein) versus the reference protein, or vice versa, as discussed more fully herein.

The determination that an antibody specifically detects a target protein or form of the target protein is made by any one of a number of standard immunoassay methods; for instance, the western blotting technique (Sambrook *et al.*, In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989). To determine that a given antibody preparation (such as one produced in a mouse) specifically detects the target protein by western blotting, total cellular protein is extracted from human cells (for example, lymphocytes) and electrophoresed on a sodium dodecyl sulfate-polyacrylamide gel. The proteins are then transferred to a membrane (for example, nitrocellulose) by western blotting, and the antibody preparation is incubated with the membrane. After washing the membrane to remove non-specifically bound antibodies, the presence of specifically bound antibodies is detected by the use of an anti-mouse antibody conjugated to an enzyme such as alkaline phosphatase.

Application of an alkaline phosphatase substrate 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium results in the production of a dense blue compound by immunolocalized alkaline phosphatase. Antibodies that specifically detect the target protein will, by this technique, be shown to bind to the target protein band (which will be localized at a given position on the gel determined by its molecular weight). Non-specific binding of the antibody to other proteins may occur and may be detectable as a weak signal on the Western blot. The non-specific nature of this binding will be recognized by one skilled in the art by the weak signal obtained on the Western blot relative to the strong primary signal arising from the specific antibody-target protein binding.

Substantially pure T2R bitter taste receptor protein or protein fragment (peptide) suitable for use as an immunogen may be isolated from the transfected or transformed cells as described above. Concentration of protein or peptide in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few micrograms per milliliter. Monoclonal or polyclonal antibody to the protein can then be prepared as follows:

A. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of the target protein identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler and Milstein (*Nature* 256:495-497, 1975) or derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody-producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess un-fused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall (*Meth. Enzymol.* 70:419-439, 1980), and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product

harvested for use. Detailed procedures for monoclonal antibody production are described in Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988).

B. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein, which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than others and may require the use of carriers and adjuvant. Also, host animals vary in response to site of inoculations and dose, with either inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appear to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis *et al.* (*J. Clin. Endocrinol. Metab.* 33:988-991, 1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony *et al.* (In *Handbook of Experimental Immunology*, Wier, D. (ed.) chapter 19. Blackwell, 1973). Plateau concentration of antibody is usually in the range of about 0.1 to 0.2 mg/ml of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher (*Manual of Clinical Immunology*, Ch. 42, 1980).

C. Antibodies Raised against Synthetic Peptides

A third approach to raising antibodies against a specific T2R bitter taste receptor protein or peptide (*e.g.*, a peptide that is specific to a variant T2R bitter taste receptor such as those disclosed herein) is to use one or more synthetic peptides synthesized on a commercially available peptide synthesizer based upon the predicted amino acid sequence of the protein or peptide. Polyclonal antibodies can be generated by injecting these peptides into, for instance, rabbits or mice.

D. Antibodies Raised by Injection of Encoding Sequence

Antibodies may be raised against proteins and peptides by subcutaneous injection of a DNA vector that expresses the desired protein or peptide, or a fragment thereof, into laboratory animals, such as mice. Delivery of the recombinant vector into the animals may be achieved using a hand-held form of the Biolistic system (Sanford *et al., Particulate Sci. Technol.* 5:27-37, 1987) as described by Tang *et al.* (*Nature* 356:152-154, 1992). Expression vectors suitable for this purpose may include those that express the T2R bitter taste receptor-encoding sequence under the transcriptional control of either the human β-actin promoter or the cytomegalovirus (CMV) promoter.

Antibody preparations prepared according to these protocols are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample; or for immunolocalization of the specified protein.

Optionally, antibodies, *e.g.*, bitter taste receptor-specific monoclonal antibodies, can be humanized by methods known in the art. Antibodies with a desired binding specificity can be commercially humanized (Scotgene, Scotland, UK; Oxford Molecular, Palo Alto, CA).

E. Antibodies Specific for Specific T2R Taste Receptor Variants

With the provision of several variant T2R bitter taste receptor proteins, the production of antibodies that specifically recognize these protein variants (and peptides derived therefrom) is enabled. In particular, production of antibodies (and fragments and engineered versions thereof) that recognize at least one variant receptor with a higher affinity than they recognize a corresponding wild type T2R bitter taste receptor, or another bitter taste receptor, is beneficial, as the resultant antibodies can be used in analysis, diagnosis and treatment (*e.g.*, inhibition or enhancement of bitter taste perception), as well as in study and examination of the T2R bitter taste receptor proteins themselves.

In particular embodiments, it is beneficial to generate antibodies from a peptide taken from a variation-specific region of the desired T2R bitter taste receptor protein. By way of example, such regions include any peptide (usually four or more

amino acids in length) that overlaps with one or more of the SNP-encoded variants described herein. More particularly, it is beneficial to raise antibodies against peptides of four or more contiguous amino acids that overlap the variants identified in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 40, 42, or 46, and particularly which comprise at least four contiguous amino acids including the residue(s) indicated in Figure 1 to be variable in different alleles if the specified T2R putative bitter taste receptors.

Longer peptides also can be used, and in some instances will produce a stronger or more reliable immunogenic response. Thus, it is contemplated in some embodiments that more than four amino acids are used to elicit the immune response, for instance, at least 5, at least 6, at least 8, at least 10, at least 12, at least 15, at least 18, at least 20, at least 25, or more, such as 30, 40, 50, or even longer peptides. Also, it will be understood by those of ordinary skill that it is beneficial in some instances to include adjuvants and other immune response enhancers, including passenger peptides or proteins, when using peptides to induce an immune response for production of antibodies.

Embodiments are not limited to antibodies that recognize epitopes containing the actual mutation identified in each variant. Instead, it is contemplated that variant-specific antibodies also may each recognize an epitope located anywhere throughout the specified T2R bitter taste receptor variant molecule, which epitopes are changed in conformation and/or availability because of the activating mutation. Antibodies directed to any of these variant-specific epitopes are also encompassed herein.

By way of example, the following references provide descriptions of methods for making antibodies specific to mutant proteins: Hills *et al.*, (*Int. J. Cancer*, 63: 537-543, 1995); Reiter & Maihle (*Nucleic Acids Res.*, 24: 4050-4056, 1996); Okamoto *et al.* (*Br. J. Cancer*, 73: 1366-1372, 1996); Nakayashiki *et al.*, (*Jpn. J. Cancer Res.*, 91: 1035-1043, 2000); Gannon *et al.* (*EMBO J.*, 9: 1595-1602, 1990); Wong *et al.* (*Cancer Res.*, 46: 6029-6033, 1986); and Carney *et al.* (*J. Cell Biochem.*, 32: 207-214, 1986). Similar methods can be employed to generate antibodies specific to specific T2R bitter taste receptor variants.

Example 8: Knockout and Overexpression Transgenic Animals

Mutant organisms that under-express or over-express one or more specific alleles of one or more specific bitter taste receptor protein are useful for research. Such mutants allow insight into the physiological and/or psychological role of bitter taste perception in a healthy and/or pathological organism. These "mutants" are "genetically engineered," meaning that information in the form of nucleotides has been transferred into the mutant's genome at a location, or in a combination, in which it would not normally exist. Nucleotides transferred in this way are said to be "non-native." For example, a non-bitter taste receptor promoter inserted upstream of a native bitter taste receptor-encoding sequence would be non-native. An extra copy of a specific bitter taste receptor gene on a plasmid, transformed into a cell, would be non-native.

Mutants may be, for example, produced from mammals, such as mice or rats, that either express, over-express, or under-express a specific allelic variant or haplotype or diplotype of a defined bitter taste receptor (or combination of bitter taste receptors), or that do not express a specified receptor (or combination of receptors) at all. Over-expression mutants are made by increasing the number of specified genes in the organism, or by introducing a specific taste receptor allele into the organism under the control of a constitutive or inducible or viral promoter such as the mouse mammary tumor virus (MMTV) promoter or the whey acidic protein (WAP) promoter or the metallothionein promoter. Mutants that under-express a taste receptor, or that do not express a specific allelic variant of a taste receptor, may be made by using an inducible or repressible promoter, or by deleting the taste receptor gene, or by destroying or limiting the function of the taste receptor gene, for instance by disrupting the gene by transposon insertion.

Antisense genes or molecules (such as siRNAs) may be engineered into the organism, under a constitutive or inducible promoter, to decrease or prevent expression of a specific T2R bitter taste receptor, as known to those of ordinary skill in the art.

A mutant mouse over-expressing a heterologous protein (such as a variant T2R bitter taste receptor protein) may be made by constructing a plasmid having a bitter taste receptor allele encoding sequence driven by a promoter, such as the

mouse mammary tumor virus (MMTV) promoter or the whey acidic protein (WAP) promoter. This plasmid may be introduced into mouse oocytes by microinjection. The oocytes are implanted into pseudopregnant females, and the litters are assayed for insertion of the transgene. Multiple strains containing the transgene are then available for study.

WAP is quite specific for mammary gland expression during lactation, and MMTV is expressed in a variety of tissues including mammary gland, salivary gland and lymphoid tissues. Many other promoters might be used to achieve various patterns of expression, *e.g.*, the metallothionein promoter.

An inducible system may be created in which the subject expression construct is driven by a promoter regulated by an agent that can be fed to the mouse, such as tetracycline. Such techniques are well known in the art.

A mutant knockout animal (*e.g.*, mouse) from which a specific taste receptor gene is deleted can be made by removing all or some of the coding regions of the gene from embryonic stem cells. The methods of creating deletion mutations by using a targeting vector have been described (Thomas and Capecchi, *Cell* 51:503-512, 1987).

Example 9: Knock-in Organisms

In addition to knock-out systems, it is also beneficial to generate "knock-ins" that have lost expression of the native protein but have gained expression of a different, usually mutant or identified allelic form of the same protein. By way of example, the allelic proteins provided herein (*e.g.*, as shown in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 34, 36, 40, 42, and 46) can be expressed in a knockout background in order to provide model systems for studying the effects of these mutants. In particular embodiments, the resultant knock-in organisms provide systems for studying taste reception, for instance how the taste of specific molecules is perceived.

Those of ordinary skill in the relevant art know methods of producing knock-in organisms. See, for instance, Rane *et al.* (*Mol. Cell Biol.*, 22: 644-656, 2002); Sotillo *et al.* (*EMBO J.*, 20: 6637-6647, 2001); Luo *et al.* (*Oncogene*, 20: 320-328, 2001); Tomasson *et al.* (*Blood*, 93: 1707-1714, 1999); Voncken *et al.* (, 86: 4603-

4611, 1995); Andrae *et al.* (*Mech. Dev.*, 107: 181-185, 2001); Reinertsen *et al.* (*Gene Expr.*, 6: 301-314, 1997); Huang *et al.* (*Mol. Med.*, 5: 129-137, 1999); Reichert *et al.* (*Blood*, 97: 1399-1403, 2001); and Huettner *et al.* (*Nat. Genet.*, 24: 57-60, 2000), by way of example.

Example 10: Screening Assays for Compounds that Modulate Taste Receptor Expression or Activity

The following assays are designed to identify compounds that interact with (*e.g.*, bind to) a variant form of a T2R bitter taste receptor (including, but not limited to an ECD or a CD or a TMD of a variant T2R bitter taste receptor), compounds that interact with (*e.g.*, bind to) intracellular proteins that interact with a variant form of a T2R bitter taste receptor (including, but not limited to, a TMD or a CD of a variant form of a T2R bitter taste receptor), compounds that interfere with the interaction of a taste receptor with transmembrane or intracellular proteins involved in taste receptor-mediated signal transduction, and to compounds which modulate the activity of a taste receptor gene (*i.e.*, modulate the level of gene expression) or modulate the level of taste receptor activity of a variant form of a T2R bitter taste receptor. Assays may additionally be utilized which identify compounds which bind to taste receptor gene regulatory sequences (*e.g.*, promoter sequences) and which may modulate taste receptor gene expression. See, *e.g.*, Platt, *J Biol Chem* 269:28558-28562, 1994.

The compounds which may be screened in accordance with the invention include, but are not limited to peptides, antibodies and fragments thereof, and other organic compounds (*e.g.*, peptidomimetics, small molecules) that bind to one or more ECDs of a variant T2R bitter taste receptor as described herein and either mimic the activity triggered by the natural ligand (*i.e.*, agonists) or inhibit the activity triggered by the natural ligand (*i.e.*, antagonists); as well as peptides, antibodies or fragments thereof, and other organic compounds that mimic the ECD of a variant T2R bitter taste receptor (or a portion thereof) and bind to and "neutralize" natural ligand.

Such compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to members of random peptide

libraries; (see, *e.g.*, Lam *et al.*, *Nature* 354:82-84, 1991; Houghten *et al.*, *Nature* 354:84-86, 1991), and combinatorial chemistry-derived molecular library made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to, members of random or partially degenerate, directed phosphopeptide libraries; see, *e.g.*, Songyang *et al.*, *Cell* 72:767-778, 1993), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, $F(ab')_2$ and Fab expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

Other compounds which can be screened in accordance with the invention include but are not limited to small organic molecules that are able to gain entry into an appropriate cell and affect the expression of a variant T2R bitter taste receptor gene or some other gene involved in a taste receptor signal transduction pathway (*e.g.*, by interacting with the regulatory region or transcription factors involved in gene expression); or such compounds that affect the activity of a variant T2R bitter taste receptor or the activity of some other intracellular factor involved in the taste receptor signal transduction pathway.

Computer modeling and searching technologies permit identification of compounds, or the improvement of already identified compounds, that can modulate expression or activity of a variant T2R bitter taste receptor. Having identified such a compound or composition, the active sites or regions are identified. Such active sites might typically be ligand binding sites, such as the interaction domains of a bitter molecule with a variant T2R bitter taste receptor itself, or the interaction domains of a bitter molecule with a specific allelic variant T2R bitter taste receptor in comparison to the interaction domains of that molecule with another variant of the same or a different T2R bitter taste receptor (to reproduce the effect of an amino acid substitution such as the alanine substitution in the PTC gene (T2R38) for designing bitter taste blockers, or to reproduce the effect of the proline substitution in the PTC gene for designing bitter taste mimics).

The active site can be identified using methods known in the art including, for example, from the amino acid sequences of peptides, from the nucleotide sequences of nucleic acids, or from study of complexes of the relevant compound or composition with its natural ligand. In the latter case, chemical methods can be used

to find the active site by finding where on the factor the complexed ligand is found. Next, the three dimensional geometric structure of the active site is determined. This can be done by known methods can determine a complete molecular structure. On the other hand, solid or liquid phase NMR can be used to determine certain intramolecular distances. Any other experimental method of structure determination can be used to obtain partial or complete geometric structures, such as high resolution electron microscopy. The geometric structures may be measured with a complexed ligand, natural or artificial, which may increase the accuracy of the active site structure determined. In another embodiment, the structure of the specified taste receptor is compared to that of a "variant" of the specified taste receptor and, rather than solve the entire structure, the structure is solved for the protein domains that are changed.

If an incomplete or insufficiently accurate structure is determined, the methods of computer based numerical modeling can be used to complete the structure or improve its accuracy. Any recognized modeling method may be used, including parameterized models specific to particular biopolymers such as proteins or nucleic acids, molecular dynamics models based on computing molecular motions, statistical mechanics models based on thermal ensembles, or combined models. For most types of models, standard molecular force fields, representing the forces between constituent atoms and groups, are necessary, and can be selected from force fields known in physical chemistry. The incomplete or less accurate experimental structures can serve as constraints on the complete and more accurate structures computed by these modeling methods.

Finally, having determined the structure of the active site, either experimentally, by modeling, or by a combination, candidate modulating compounds can be identified by searching databases containing compounds along with information on their molecular structure. Such a search seeks compounds having structures that match the determined active site structure and that interact with the groups defining the active site. Such a search can be manual, but is preferably computer assisted. These compounds found from this search are potential variant T2R bitter taste receptor modulating compounds.

Alternatively, these methods can be used to identify improved modulating compounds from an already known modulating compound or ligand. The composition of the known compound can be modified and the structural effects of modification can be determined using the experimental and computer modeling methods described above applied to the new composition. The altered structure is then compared to the active site structure of the compound to determine if an improved fit or interaction results. In this manner systematic variations in composition, such as by varying side groups, can be quickly evaluated to obtain modified modulating compounds or ligands of improved specificity or activity.

In another embodiment, the structure of a specified allelic taste receptor (the reference form) is compared to that of a variant taste receptor (encoded by a different allele of the same specified receptor). Then, potential bitter taste inhibitors are designed that bring about a structural change in the reference form so that it resembles the variant form. Or, potential bitter taste mimics are designed that bring about a structural change in the variant form so that it resembles another variant form, or the form of the reference receptor.

Further experimental and computer modeling methods useful to identify modulating compounds based upon identification of the active sites of bitter compounds, various variants of the T2R bitter taste receptors described herein, and related transduction and transcription factors will be apparent to those of skill in the art.

Examples of molecular modeling systems are the CHARMM and QUANTA programs (Polygen Corporation, Waltham, Mass.). CHARMm performs the energy minimization and molecular dynamics functions. QUANTA performs the construction, graphic modeling and analysis of molecular structure. QUANTA allows interactive construction, modification, visualization, and analysis of the behavior of molecules with each other.

A number of articles review computer modeling of drugs interactive with specific-proteins, such as Rotivinen *et al. Acta Pharmaceutical Fennica* 97:159-166, 1988; Ripka, *New Scientist* 54-57, 1988; McKinaly and Rossmann, *Annu Rev Pharmacol Toxicol* 29:111-122, 1989; Perry and Davies, *OSAR: Quantitative Structure-Activity Relationships in Drug Design* pp. 189-193, 1989 (Alan R. Liss,

Inc.); Lewis and Dean, *Proc R Soc Lond* 236:125-140 and 141-162, 1989; and, with respect to a model receptor for nucleic acid components, Askew *et al.*, *J Am Chem Soc* 111:1082-1090, 1989. Other computer programs that screen and graphically depict chemicals are available from companies such as BioDesign, Inc. (Pasadena, Calif.), Allelix, Inc. (Mississauga, Ontario, Canada), and Hypercube, Inc. (Cambridge, Ontario). Although these are primarily designed for application to drugs specific to particular proteins, they can be adapted to design of drugs specific to regions of DNA or RNA, once that region is identified.

Although described above with reference to design and generation of compounds which could alter binding, one could also screen libraries of known compounds, including natural products or synthetic chemicals, and biologically active materials, including proteins, for compounds which are inhibitors or activators.

Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of a variant T2R bitter taste receptor gene product, and for designing bitter taste blockers and mimics.

Example 11: In vitro Screening Assays for Compounds that Bind to a Variant T2R Taste Receptor

In vitro systems may be designed to identify compounds capable of interacting with (*e.g.*, binding to) a variant T2R bitter taste receptor (including, but not limited to, an ECD, or a TMD, or a CD of a variant T2R bitter taste receptor). Compounds identified may be useful, for example, in modulating the activity of "wild type" and/or "variant" T2R bitter taste receptor gene products; may be useful in elaborating the biological function of taste receptors; may be utilized in screens for identifying compounds that disrupt normal T2R bitter taste receptor interactions; or may in themselves disrupt such interactions.

The principle of assays used to identify compounds that bind to a variant T2R bitter taste receptor involves preparing a reaction mixture of a variant T2R bitter taste receptor polypeptide and a test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex which can be removed and/or detected in the reaction mixture. The variant T2R

bitter taste receptor species used can vary depending upon the goal of the screening assay. For example, where agonists or antagonists are sought, the full length variant T2R bitter taste receptor, or a soluble truncated taste receptor, *e.g.*, in which a TMD and/or a CD is deleted from the molecule, a peptide corresponding to an ECD or a fusion protein containing a variant T2R bitter taste receptor ECD fused to a protein or polypeptide that affords advantages in the assay system (*e.g.*, labeling, isolation of the resulting complex, etc.) can be utilized. Where compounds that interact with the cytoplasmic domain are sought to be identified, peptides corresponding to a variant T2R bitter taste receptor CD and fusion proteins containing a variant T2R bitter taste receptor CD can be used.

The screening assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring the variant T2R bitter taste receptor protein, polypeptide, peptide or fusion protein or the test substance onto a solid phase and detecting taste receptor/test compound complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the taste receptor reactant may be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtiter plates may conveniently be utilized as the solid phase. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect

label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for a variant T2R bitter taste receptor protein, polypeptide, peptide or fusion protein or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

Alternatively, cell-based assays, membrane vesicle-based assays and membrane fraction-based assays can be used to identify compounds that interact with a variant T2R bitter taste receptor. To this end, cell lines that express a variant T2R bitter taste receptor (or combination thereof) or cell lines (*e.g.*, COS cells, CHO cells, HEK293 cells, etc.) have been genetically engineered to express variant T2R bitter taste receptor (*e.g.*, by transfection or transduction of taste receptor DNA) can be used. Interaction of the test compound with, for example, an ECD or a CD of a variant T2R bitter taste receptor expressed by the host cell can be determined by comparison or competition with a bitter compound or analog thereof, such as PTC.

A variant T2R bitter taste receptor polypeptide (such as those described herein) may be employed in a screening process for compounds which bind the receptor and which activate (agonists) or inhibit activation (antagonists) of the receptor polypeptide of the present invention. Thus, polypeptides described herein may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan *et al. Current Protocols in Immunology* 1 (2): Chapter 5, 1991.

In general, such screening procedures involve providing appropriate cells which express a receptor polypeptide of the present invention on the surface thereof. Such cells include cells from mammals, insects, yeast, and bacteria. In particular, a polynucleotide encoding the receptor of the present invention is employed to transfect cells to thereby express a variant T2R bitter taste receptor. The expressed

receptor is then contacted with a test compound to observe binding, stimulation or inhibition of a functional response.

One such screening procedure involves the use of melanophores that are transfected to express a variant T2R bitter taste receptor. Such a screening technique is described in PCT WO 92/01810, published Feb. 6, 1992. Such an assay may be employed to screen for a compound which inhibits activation of a receptor of the present invention by contacting the melanophore cells which encode the receptor with both a receptor ligand, such as PTC or another bitter compound, and a compound to be screened. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor, *i.e.*, inhibits activation of the receptor.

The technique may also be employed for screening of compounds which activate a receptor of the present invention by contacting such cells with compounds to be screened and determining whether such compound generates a signal, *i.e.*, activates the receptor.

Other screening techniques include the use of cells which express a variant T2R bitter taste receptor (for example, transfected CHO cells) in a system which measures extracellular pH changes caused by receptor activation. In this technique, compounds may be contacted with cells expressing a receptor polypeptide of the present invention. A second messenger response, *e.g.*, signal transduction or pH changes, is then measured to determine whether the potential compound activates or inhibits the receptor.

Another screening technique involves expressing a variant T2R bitter taste receptor in which the receptor is linked to phospholipase C or D. Representative examples of such cells include, but are not limited to, endothelial cells, smooth muscle cells, and embryonic kidney cells. The screening may be accomplished as hereinabove described by detecting activation of the receptor or inhibition of activation of the receptor from the phospholipase second signal.

Another method involves screening for compounds which are antagonists, and thus inhibit activation of a receptor polypeptide of the present invention by determining inhibition of binding of labeled ligand, such as PTC or another bitter compound, to cells which have the receptor on the surface thereof, or cell

membranes containing the receptor. Such a method involves transfecting a eukaryotic cell with a DNA encoding a variant T2R bitter taste receptor such that the cell expresses the receptor on its surface, or using of eukaryotic cells that express the receptor of the present invention on their surface (or using a eukaryotic cell that expresses the receptor on its surface). The cell is then contacted with a potential antagonist in the presence of a labeled form of a ligand, such as PTC or another bitter comound. The ligand can be labeled, *e.g.*, by radioactivity. The amount of labeled ligand bound to the receptors is measured, *e.g.*, by measuring radioactivity associated with transfected cells or membrane from these cells. If the compound binds to the receptor, the binding of labeled ligand to the receptor is inhibited as determined by a reduction of labeled ligand that binds to the receptors. This method is called a binding assay.

Another such screening procedure involves the use of eukaryotic cells, which are transfected to express the receptor of the present invention, or use of eukaryotic cells that express the receptor of the present invention on their surface. The cells are loaded with an indicator dye that produces a fluorescent signal when bound to calcium, and the cells are contacted with a test substance and a receptor agonist, such as PTC or another bitter compound. Any change in fluorescent signal is measured over a defined period of time using, for example, a fluorescence spectrophotometer or a fluorescence imaging plate reader. A change in the fluorescence signal pattern generated by the ligand indicates that a compound is a potential antagonist (or agonist) for the receptor.

Another such screening procedure involves use of eukaryotic cells, which are transfected to express the receptor of the present invention (or use of eukaryotic cells that express the receptor of the present invention), and which are also transfected with a reporter gene construct that is coupled to activation of the receptor (for example, luciferase or beta-galactosidase behind an appropriate promoter). The cells are contacted with a test substance and a receptor agonist, such as PTC or another bitter compound, and the signal produced by the reporter gene is measured after a defined period of time. The signal can be measured using a luminometer, spectrophotometer, fluorimeter, or other such instrument appropriate for the specific

reporter construct used. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor.

Another such screening technique for antagonists or agonists involves introducing RNA encoding a PTC taste receptor into *Xenopus* oocytes to transiently express the receptor. The receptor expressing oocytes are then contacted with a receptor ligand, such as PTC, and a compound to be screened. Inhibition or activation of the receptor is then determined by detection of a signal, such as, cAMP, calcium, proton, or other ions.

Another such technique of screening for antagonists or agonists involves determining inhibition or stimulation of T2R taste receptor-mediated cAMP and/or adenylate cyclase accumulation or diminution. Such a method involves transiently or stably transfecting a eukaryotic cell with a variant T2R bitter taste receptor to express the receptor on the cell surface (or using a eukaryotic cell that expresses the receptor of the present invention on its surface). The cell is then exposed to potential antagonists in the presence of ligand, such as PTC or another bitter compound. The amount of cAMP accumulation is then measured, for example, by radio-immuno or protein binding assays (for example using Flashplates or a scintillation proximity assay). Changes in cAMP levels can also be determined by directly measuring the activity of the enzyme, adenylyl cyclase, in broken cell preparations. If the potential antagonist binds the receptor, and thus inhibits taste receptor binding, the levels of variant T2R bitter taste receptor-mediated cAMP, or adenylate cyclase activity, will be reduced or increased.

Example 12: Assays for Intracellular Proteins that Interact with a Variant T2R Bitter Taste Receptor.

Any method suitable for detecting protein-protein interactions may be employed for identifying transmembrane proteins or intracellular proteins that interact with a variant T2R bitter taste receptor. Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns of cell lysates or proteins obtained from cell lysates and a variant T2R bitter taste receptor to identify proteins in the lysate that interact with the PTC taste receptor. For these assays, a variant T2R

bitter taste receptor component used can be a full length taste receptor, a soluble derivative lacking the membrane-anchoring region (*e.g.*, a truncated taste receptor in which all TMDs are deleted resulting in a truncated molecule containing ECDs fused to CDs), a peptide corresponding to a CD or a fusion protein containing a CD of PTC taste receptor.

Once isolated, such an intracellular protein can be identified and can, in turn, be used, in conjunction with standard techniques, to identify proteins with which it interacts. For example, at least a portion of the amino acid sequence of an intracellular protein which interacts with the variant T2R bitter taste receptor can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique. See, *e.g.*, Creighton *Proteins: Structures and Molecular Principles*, W.H. Freeman & Co., N.Y., pp. 34-49, 1983. The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding such intracellular proteins. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well known. See, *e.g.*, Ausubel *et al. Current Protocols in Molecular Biology* Green Publishing Associates and Wiley Interscience, N.Y., 1989; and Innis *et al.*, eds. *PCR Protocols: A Guide to Methods and Applications* Academic Press, Inc., New York, 1990.

Additionally, methods may be employed which result in the simultaneous identification of genes, which encode the transmembrane or intracellular proteins interacting with a variant T2R bitter taste receptor. These methods include, for example, probing expression libraries, in a manner similar to the well known technique of antibody probing of λgt11 libraries, using labeled PTC taste receptor protein, or a variant T2R bitter taste receptor polypeptide, peptide or fusion protein, *e.g.*, a variant T2R bitter taste receptor polypeptide or PTC taste receptor domain fused to a marker (*e.g.*, an enzyme, fluor, luminescent protein, or dye), or an Ig-Fc domain.

One method that detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version

of this system has been described (Chien *et al.*, *PNAS USA* 88:9578-9582, 1991) and is commercially available from Clontech (Palo Alto, Calif.).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one plasmid consists of nucleotides encoding the DNA-binding domain of a transcription activator protein fused to a variant T2R bitter taste receptor nucleotide sequence encoding a variant T2R bitter taste receptor, a variant T2R bitter taste receptor polypeptide, peptide or fusion protein, and the other plasmid consists of nucleotides encoding the transcription activator protein's activation domain fused to a cDNA encoding an unknown protein which has been recombined into this plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (*e.g.*, HBS or lacZ) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with the "bait" gene product. By way of example, and not by way of limitation, PTC taste receptor may be used as the bait gene product. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of a bait variant T2R bitter taste receptor gene product fused to the DNA-binding domain are cotransformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, and not by way of limitation, a bait variant T2R bitter taste receptor gene sequence, such as the open reading frame of variant T2R bitter taste receptor (or a domain of a taste receptor) can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the

library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait variant T2R bitter taste receptor gene product are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments can be inserted into a vector such that they are translationally fused to the transcriptional activation domain of GAL4. This library can be co-transformed along with the bait PTC taste receptor gene-GAL4 fusion plasmid into a yeast strain, which contains a lacZ gene driven by a promoter that contains GAL4 activation sequence. A cDNA encoded protein, fused to GAL4 transcriptional activation domain, that interacts with bait PTC taste receptor gene product will reconstitute an active GAL4 protein and thereby drive expression of the HIS3 gene. Colonies, which express HIS3, can be detected by their growth on Petri dishes containing semi-solid agar based media lacking histidine. The cDNA can then be purified from these strains, and used to produce and isolate the bait PTC taste receptor gene-interacting protein using techniques routinely practiced in the art.

Example 13: Assays for Compounds that Interfere with Taste Receptor /Intracellular or Taste Receptor /Transmembrane Macromolecule Interaction

The macromolecules that interact with a variant T2R bitter taste receptor are referred to, for purposes of this discussion, as "binding partners". These binding partners are likely to be involved in a variant T2R bitter taste receptor signal transduction pathway, and therefore, in the role of taste receptors and taste receptor variants in bitter tasting. Therefore, it is desirable to identify compounds that interfere with or disrupt the interaction of such binding partners with variant and/or normal T2R bitter taste receptor, which may be useful in regulating the activity of variant T2R bitter taste receptors and control the sensitivity to bitter tastes associated with certain taste receptor activity.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between a variant T2R bitter taste receptor and its binding partner or partners involves preparing a reaction mixture containing variant T2R bitter taste receptor protein, polypeptide, peptide or fusion protein as described

above, and the binding partner under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of a variant T2R bitter taste receptor moiety and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between a variant T2R bitter taste receptor moiety and the binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of a variant T2R bitter taste receptor and the binding partner. Additionally, complex formation within reaction mixtures containing the test compound and reference T2R bitter taste receptor variant may also be compared to complex formation within reaction mixtures containing the test compound and a different allelic or other variant of the same T2R taste receptor. This comparison may be important in those cases wherein it is desirable to identify compounds that disrupt interactions of reference but not variant T2R taste receptors, or differentially disrupt interactions between different variant T2R taste receptors.

The assay for compounds that interfere with the interaction of a variant T2R bitter taste receptor and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either a variant T2R bitter taste receptor moiety product or the binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction by competition can be identified by conducting the reaction in the presence of the test substance; *i.e.*, by adding the test substance to the reaction mixture prior to or simultaneously with a variant T2R bitter taste receptor moiety and interactive binding partner. Alternatively, test compounds that disrupt preformed complexes, *e.g.*, compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound

to the reaction mixture after complexes have been formed. The various formats are described briefly below.

In a heterogeneous assay system, either a variant T2R bitter taste receptor moiety or the interactive binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of a variant T2R bitter taste receptor gene product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex or which disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of a variant T2R bitter taste receptor moiety and the interactive binding partner is prepared in which either a variant T2R bitter taste receptor or its binding partners is labeled, but the signal generated by the label is quenched due to formation of the complex (see, *e.g.*, U.S. Pat. No. 4,109,496 by Rubenstein, which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances, which disrupt PTC taste receptor/intracellular binding partner interaction can be identified.

In a particular embodiment, a variant T2R bitter taste receptor fusion can be prepared for immobilization. For example, a variant T2R bitter taste receptor or a peptide fragment, *e.g.*, corresponding to a CD, can be fused to a glutathione-S-transferase (GST) gene using a fusion vector, such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion protein. The interactive binding partner can be purified and used to raise a monoclonal antibody, using methods routinely practiced in the art and described above. This antibody can be labeled with the radioactive isotope ^{125}I, for example, by methods routinely practiced in the art. In a heterogeneous assay, *e.g.*, the GST-taste receptor fusion protein can be anchored to glutathione-agarose beads. The interactive binding partner can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the complexed components. The interaction between a variant T2R bitter taste receptor gene product and the interactive binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

Alternatively, the GST-taste receptor fusion protein and the interactive binding partner can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the species are allowed to interact. This mixture can then be added to the

glutathione-agarose beads and unbound material is washed away. Again the extent of inhibition of a variant T2R bitter taste receptor/binding partner interaction can be detected by adding the labeled antibody and measuring the radioactivity associated with the beads.

In another embodiment, these same techniques can be employed using peptide fragments that correspond to the binding domains of a variant T2R bitter taste receptor and/or the interactive or binding partner (in cases where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not limited to, mutagenesis of the gene encoding one of the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond to the region of the protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the intracellular binding partner is obtained, short gene segments can be engineered to express peptide fragments of the protein, which can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, a variant T2R bitter taste receptor gene product can be anchored to a solid material as described, above, by making a GST-taste receptor fusion protein and allowing it to bind to glutathione agarose beads. The interactive binding partner can be labeled with a radioactive isotope, such as ^{35}S, and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-taste receptor fusion protein and allowed to bind. After washing away unbound peptides, labeled bound material, representing the intracellular binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides

so identified can be produced synthetically or fused to appropriate facilitative proteins using recombinant DNA technology.

Example 14: Assays for Identification of Compounds that Modulate Bitter Tastes

Compounds, including but not limited to compounds identified via assay techniques such as those described above, can be tested for the ability to modulate bitter tastes. The assays described above can identify compounds that affect variant T2R bitter taste receptor activity (*e.g.*, compounds that bind to a variant T2R bitter taste receptor, inhibit binding of the natural ligand, and either activate signal transduction (agonists) or block activation (antagonists), and compounds that bind to a ligand of a variant T2R bitter taste receptor and neutralize ligand activity); or compounds that affect variant T2R bitter taste receptor gene activity (by affecting T2R bitter taste receptor gene expression, including molecules, *e.g.*, proteins or small organic molecules, that affect or interfere with events so that expression of the full length variant or wild-tyep T2R bitter taste receptor can be modulated). However, it should be noted that the assays described can also identify compounds that modulate variant T2R bitter taste receptor signal transduction (*e.g.*, compounds which affect downstream signaling events, such as inhibitors or enhancers of protein kinases or phosphatases activities which participate in transducing the signal activated by binding of a bitter compound (*e.g.*, PTC) to a variant T2R bitter taste receptor). The identification and use of such compounds which affect another step in a variant T2R bitter taste receptor signal transduction pathway in which a variant T2R bitter taste receptor and/or variant T2R bitter taste receptor gene product is involved and, by affecting this same pathway may modulate the effect of variant T2R bitter taste receptor on the sensitivity to bitter tastes are within the scope of the invention. Such compounds can be used as part of a therapeutic method for modulating bitter tastes.

Cell-based systems, membrane vesicle-based systems and membrane fraction-based systems can be used to identify compounds that may act to modulate bitter tastes. Such cell systems can include, for example, recombinant or non-recombinant cells, such as cell lines, which express the PTC taste receptor gene. In addition, expression host cells (*e.g.*, COS cells, CHO cells, HEK293 cells)

genetically engineered to express a functional variant T2R bitter taste receptor and to respond to activation by the natural ligand, *e.g.*, as measured by a chemical or phenotypic change, induction of another host cell gene, change in ion flux (*e.g.*, Ca^{2+}), phosphorylation of host cell proteins, etc., can be used as an end point in the assay.

In utilizing such cell systems, cells may be exposed to a compound suspected of exhibiting an ability to modulate bitter tastes, at a sufficient concentration and for a time sufficient to elicit such a modulation in the exposed cells. After exposure, the cells can be assayed to measure alterations in the expression of a variant T2R bitter taste receptor gene, *e.g.*, by assaying cell lysates for PTC taste receptor mRNA transcripts (*e.g.*, by Northern analysis) or for variant T2R bitter taste receptor protein expressed in the cell; compounds which regulate or modulate expression of a variant T2R bitter taste receptor gene are good candidates as therapeutics. Alternatively, the cells are examined to determine whether one or more cellular phenotypes has been altered to resemble a taster or nontaster type. Still further, the expression and/or activity of components of the signal transduction pathway of which a variant T2R bitter taste receptor is a part, or the activity of a T2R bitter taste receptor signal transduction pathway itself can be assayed.

For example, after exposure, the cell lysates can be assayed for the presence of phosphorylation of host cell proteins, as compared to lysates derived from unexposed control cells. The ability of a test compound to inhibit phosphorylation of host cell proteins in these assay systems indicates that the test compound alters signal transduction initiated by taste receptor activation. The cell lysates can be readily assayed using a Western blot format; *i.e.*, the host cell proteins are resolved by gel electrophoresis, transferred and probed using a detection antibody (*e.g.*, an antibody labeled with a signal generating compound, such as radiolabel, fluor, enzyme, etc.), see, *e.g.*, Glenney *et al.*, *J Immunol Methods* 109:277-285, 1988; Frackelton *et al.*, *Mol Cell Biol* 3:1343-1352, 1983. Alternatively, an ELISA format could be used in which a particular host cell protein involved in the taste receptor signal transduction pathway is immobilized using an anchoring antibody specific for the target host cell protein, and the presence or absence of a phosphorylated residue on the immobilized host cell protein is detected using a labeled antibody. (See, *e.g.*,

King *et al.*, *Life Sci* 53:1465-1472, 1993). In yet another approach, ion flux, such as calcium ion flux, can be measured as an end point for PTC taste receptor stimulated signal transduction.

Example 15: Other Assays for Modulators of Variant T2R Bitter Taste Receptors

A. Assays for Taste Receptor Protein Activity

T2R bitter taste receptor family members are G-protein coupled receptors that participate in taste transduction, *e.g.*, bitter taste transduction. The activity of a T2R bitter taste receptor protein variants can be assessed using a variety of *in vitro* and *in vivo* assays to determine functional, chemical, and physical effects, *e.g.*, measuring ligand binding (*e.g.*, radioactive ligand binding), second messengers (*e.g.*, cAMP, cGMP, IP_3, DAG, or Ca^{2+}), ion flux, phosphorylation levels, transcription levels, neurotransmitter levels, and the like. Furthermore, such assays can be used to test for inhibitors and activators of identified T2R bitter taste receptor family member variants. Modulators can also be genetically altered versions of taste receptors. Such modulators of taste transduction activity are useful for customizing taste, for example to modify the detection of bitter tastes.

Modulators of a T2R bitter taste receptor protein variant activity are tested using taste receptor polypeptides as described herein, either recombinant or naturally occurring. The protein can be isolated, expressed in a cell, expressed in a membrane derived from a cell, expressed in tissue or in an animal, either recombinant or naturally occurring. For example, tongue slices, dissociated cells from a tongue, transformed cells, or membranes can be used. Modulation is tested using one of the *in vitro* or *in vivo* assays described herein. Taste transduction can also be examined *in vitro* with soluble or solid state reactions, using a full-length taste receptor or a chimeric molecule such as an extracellular domain or transmembrane domain, or combination thereof, of a taste receptor variant covalently linked to a heterologous signal transduction domain, or a heterologous extracellular domain and/or transmembrane domain covalently linked to the transmembrane and/or cytoplasmic domain of a T2R bitter taste receptor protein variant. Furthermore, ligand-binding domains of the protein of interest can be used *in vitro* in soluble or solid state reactions to assay for ligand binding. In numerous embodiments, a chimeric

receptor will be made that comprises all or part of a T2R bitter taste receptor protein variant as well an additional sequence that facilitates the localization of the taste receptor to the membrane, such as a rhodopsin, *e.g.*, an N-terminal fragment of a rhodopsin protein.

Ligand binding a T2R bitter taste receptor protein variant, a domain, or chimeric protein can be tested in solution, in a bilayer membrane, attached to a solid phase, in a lipid monolayer, or in vesicles. Binding of a modulator can be tested using, *e.g.*, changes in spectroscopic characteristics (*e.g.*, fluorescence, absorbance, refractive index) hydrodynamic (*e.g.*, shape), chromatographic, or solubility properties.

Receptor-G-protein interactions can also be examined. For example, binding of the G-protein to the receptor or its release from the receptor can be examined. For example, in the absence of GTP, an activator will lead to the formation of a tight complex of a G protein (all three known subunits) with the receptor. This complex can be detected in a variety of ways, as noted above. Such an assay can be modified to search for inhibitors, *e.g.*, by adding an activator to the receptor and G protein in the absence of GTP, which form a tight complex, and then screen for inhibitors by looking at dissociation of the receptor-G protein complex. In the presence of GTP, release of the known alpha subunit of the G protein from the other two known G protein subunits serves as a criterion of activation.

In a convenient embodiment, T2R bitter taste receptor protein variant-gustducin interactions are monitored as a function of taste receptor activation. One taste-cell specific G protein that has been identified is called gustducin (McLaughlin *et al. Nature* 357:563-569, 1992). Such ligand dependent coupling of taste receptors with gustducin can be used as a marker to identify modifiers of the T2R bitter taste receptor protein variant.

An activated or inhibited G-protein will in turn alter the properties of target enzymes, channels, and other effector proteins. The classic examples are the activation of cGMP phosphodiesterase by transducin in the visual system, adenylate cyclase by the stimulatory G-protein, phospholipase C by Gq and other cognate G proteins, and modulation of diverse channels by Gi and other G proteins.

Downstream consequences can also be examined such as generation of diacyl glycerol and IP3 by phospholipase C, and in turn, for calcium mobilization by IP3.

In a convenient embodiment, a T2R bitter taste receptor protein variant is expressed in a eukaryotic cell as a chimeric receptor with a heterologous, chaperone sequence that facilitates its maturation and targeting through the secretory pathway. In a preferred embodiment, the heterologous sequence is a rhodopsin sequence, such as an N-terminal leader of a rhodopsin. Such chimeric taste receptors can be expressed in any eukaryotic cell, such as HEK293 cells. Preferably, the cells comprise a functional G protein, *e.g.*, Gα15, that is capable of coupling the chimeric receptor to an intracellular signaling pathway or to a signaling protein such as phospholipase Cβ. Activation of such chimeric receptors in such cells can be detected using any standard method, such as by detecting changes in intracellular calcium by detecting FURA-2 dependent fluorescence in the cell.

An activated G-protein coupled receptor (GPCR) becomes a substrate for kinases that phosphorylate the C-terminal tail of the receptor (and possibly other sites as well). Thus, activators will promote the transfer of ^{32}P from gamma-labeled GTP to the receptor, which can be assayed with a scintillation counter. The phosphorylation of the C-terminal tail will promote the binding of arrestin-like proteins and will interfere with the binding of G-proteins. The kinase/arrestin pathway plays a key role in the desensitization of many GPCR receptors. For example, compounds that modulate the duration a taste receptor stays active would be useful as a means of prolonging a desired taste or cutting off an unpleasant one. For a general review of GPCR signal transduction and methods of assaying signal transduction, see, *e.g.*, *Methods in Enzymology*, vols. 237 and 238 (1994) and volume 96 (1983); Bourne *et al.*, *Nature* 10:349:117-127, 1991; Bourne *et al.*, *Nature* 348:125-132, 1990; Pitcher *et al.*, *Annu Rev Biochem* 67:653-692, 1998.

Samples or assays that are treated with a potential T2R bitter taste receptor protein variant inhibitor or activator are compared to control samples without the test compound, to examine the extent of modulation. Such assays may be carried out in the presence of a bitter tastant that is known to activate the particular receptor, and modulation of the bitter-tastant-dependent activation monitored. Control samples (untreated with activators or inhibitors) are assigned a relative T2R bitter

taste receptor protein activity value of 100. Inhibition of a T2R bitter taste receptor protein variant is achieved when the T2R bitter taste receptor protein variant activity value relative to the control is about 90%, optionally 50%, optionally 25-0%. Activation of a T2R bitter taste receptor protein variant is achieved when the T2R bitter taste receptor protein variant activity value relative to the control is 110%, optionally 150%, 200-500%, or 1000-2000%.

Changes in ion flux may be assessed by determining changes in polarization (*i.e.*, electrical potential) of the cell or membrane expressing a T2R bitter taste receptor protein variant. One means to determine changes in cellular polarization is by measuring changes in current (thereby measuring changes in polarization) with voltage-clamp and patch-clamp techniques, *e.g.*, the "cell-attached" mode, the "inside-out" mode, and the "whole cell" mode (see, *e.g.*, Ackerman *et al., New Engl J Med* 336:1575-1595, 1997). Whole cell currents are conveniently determined using the standard methodology (see, *e.g.*, Hamil *et al., Pflugers Archiv* 391:85, 1981). Other known assays include: radiolabeled ion flux assays and fluorescence assays using voltage-sensitive dyes (see, *e.g.*, Vestergarrd-Bogind *et al., J Membrane Biol* 88:67-75, 1988; Gonzales & Tsien, *Chem Biol* 4:269-277, 1997; Daniel *et al., J Pharmacol Meth* 25:185-193, 1991; Holevinsky *et al., J Membrane Biology* 137:59-70, 1994). Generally, the compounds to be tested are present in the range from 1 pM to 100 mM.

The effects of the test compounds upon the function of the polypeptides can be measured by examining any of the parameters described above. Any suitable physiological change that affects GPCR activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as transmitter release, hormone release, transcriptional changes to both known and uncharacterized genetic markers (*e.g.*, northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as Ca^{2+}, IP3, cGMP, or cAMP.

Convenient assays for G-protein coupled receptors include cells that are loaded with ion or voltage sensitive dyes to report receptor activity. Assays for determining activity of such receptors can also use known agonists and antagonists

for other G-protein coupled receptors as negative or positive controls to assess activity of tested compounds. In assays for identifying modulatory compounds (*e.g.*, agonists, antagonists), changes in the level of ions in the cytoplasm or membrane voltage will be monitored using an ion sensitive or membrane voltage fluorescent indicator, respectively. Among the ion-sensitive indicators and voltage probes that may be employed are those disclosed in the Molecular Probes 1997 Catalog. For G-protein coupled receptors, promiscuous G-proteins such as Gα15 and Gα16 can be used in the assay of choice (Wilkie *et al. PNAS USA* 88:10049-10053, 1991). Such promiscuous G-proteins allow coupling of a wide range of receptors.

Receptor activation typically initiates subsequent intracellular events, *e.g.*, increases in second messengers such as IP3, which releases intracellular stores of calcium ions. Activation of some G-protein coupled receptors stimulates the formation of inositol triphosphate (IP3) through phospholipase C-mediated hydrolysis of phosphatidylinositol (Berridge & Irvine *Nature*, 312:315-321, 1984). IP3 in turn stimulates the release of intracellular calcium ion stores. Thus, a change in cytoplasmic calcium ion levels, or a change in second messenger levels such as IP3 can be used to assess G-protein coupled receptor function. Cells expressing such G-protein coupled receptors may exhibit increased cytoplasmic calcium levels as a result of contribution from both intracellular stores and via activation of ion channels, in which case it may be desirable although not necessary to conduct such assays in calcium-free buffer, optionally supplemented with a chelating agent such as EGTA, to distinguish fluorescence response resulting from calcium release from internal stores.

Other assays can involve determining the activity of receptors which, when activated, result in a change in the level of intracellular cyclic nucleotides, *e.g.*, cAMP or cGMP, by activating or inhibiting enzymes such as adenylate cyclase. There are cyclic nucleotide-gated ion channels, *e.g.*, rod photoreceptor cell channels and olfactory neuron channels that are permeable to cations upon activation by binding of cAMP or cGMP (see, *e.g.*, Altenhofen *et al., PNAS USA* 88:9868-9872, 1991; and Dhallan *et al., Nature* 347:184-187, 1990). In cases where activation of the receptor results in a decrease in cyclic nucleotide levels, it may be preferable to expose the cells to agents that increase intracellular cyclic nucleotide levels, *e.g.*,

forskolin, prior to adding a receptor-activating compound to the cells in the assay. Cells for this type of assay can be made by co-transfection of a host cell with DNA encoding a cyclic nucleotide-crated ion channel, GPCR phosphatase and DNA encoding a receptor (*e.g.*, certain glutamate receptors, muscarinic acetylcholine receptors, dopamine receptors, serotonin receptors, and the like), which, when activated, causes a change in cyclic nucleotide levels in the cytoplasm.

In a convenient embodiment, a T2R bitter taste receptor protein variant activity is measured by expressing a T2R bitter taste receptor protein variant gene in a heterologous cell with a promiscuous G-protein that links the receptor to a phospholipase C signal transduction pathway (see, Offermanns & Simon, *J Biol Chem* 270:15175-15180, 1995). Optionally the cell line is HEK293 (which does not naturally express PTC taste receptor genes and the promiscuous G-protein is Gα15 (Offermanns & Simon, 1995). Modulation of taste transduction is assayed by measuring changes in intracellular Ca^{2+} levels, which change in response to modulation of the a T2R bitter taste receptor protein variant signal transduction pathway via administration of a molecule that associates with a a T2R bitter taste receptor protein variant. Changes in Ca^{2+} levels are optionally measured using fluorescent Ca^{2+} indicator dyes and fluorometric imaging.

In one embodiment, the changes in intracellular cAMP or cGMP can be measured using immunoassays. The method described in Offermanns & Simon (*J Biol Chem* 270:15175-15180, 1995), for instance, may be used to determine the level of cAMP. Also, the method described in Felley-Bosco *et al.* (*Am J Resp Cell and Mol Biol* 11:159-164, 1994) may be used to determine the level of cGMP. Further, an assay kit for measuring cAMP and/or cGMP is described in U.S. Patent 4,115,538.

In another embodiment, phosphatidyl inositol (PI) hydrolysis can be analyzed according to U.S. Patent 5,436,128. Briefly, the assay involves labeling of cells with ^{3}H-myoinositol for 48 or more hours. The labeled cells are treated with a test compound for one hour. The treated cells are lysed and extracted in chloroform-methanol-water after which the inositol phosphates are separated by ion exchange chromatography and quantified by scintillation counting. Fold stimulation is determined by calculating the ratio of cpm in the presence of agonist to cpm in the

presence of buffer control. Likewise, fold inhibition is determined by calculating the ratio of cpm in the presence of antagonist to cpm in the presence of buffer control (which may or may not contain an agonist).

In another embodiment, transcription levels can be measured to assess the effects of a test compound on signal transduction. A host cell containing a T2R bitter taste receptor protein variant of interest is contacted with a test compound for a sufficient time to effect any interactions, and then the level of gene expression is measured. The amount of time to effect such interactions may be empirically determined, such as by running a time course and measuring the level of transcription as a function of time. The amount of transcription may be measured by using any method known to those of skill in the art to be suitable. For example, mRNA expression of the protein of interest may be detected using northern blots or their polypeptide products may be identified using immunoassays. Alternatively, transcription based assays using reporter genes may be used as described in U.S. Patent 5,436,128. The reporter genes can be, *e.g.*, chloramphenicol acetyltransferase, luciferase, β-galactosidase and alkaline phosphatase. Furthermore, the protein of interest can be used as an indirect reporter via attachment to a second reporter such as green fluorescent protein (see, *e.g.*, Mistili & Spector, *Nature Biotechnology* 15:961-964, 1997).

The amount of transcription is then compared to the amount of transcription in either the same cell in the absence of the test compound, or it may be compared with the amount of transcription in a substantially identical cell that lacks the protein of interest. A substantially identical cell may be derived from the same cells from which the recombinant cell was prepared but which had not been modified by introduction of heterologous DNA. Any difference in the amount of transcription indicates that the test compound has in some manner altered the activity of the protein of interest.

B. Modulators

The compounds tested as modulators of a T2R bitter taste receptor family member variant can be any small chemical compound, or a biological entity, such as a protein, sugar, nucleic acid or lipid. Alternatively, modulators can be genetically

altered versions of a T2R bitter taste receptor protein gene. Typically, test compounds will be small chemical molecules and peptides. Essentially any chemical compound can be used as a potential modulator or ligand in the assays of the invention, although most often compounds dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (*e.g.*, in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs, Switzerland) and the like.

In one convenient embodiment, high throughput screening methods involve providing a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such "combinatorial chemical libraries" or "ligand libraries" are then screened in one or more assays, as described herein, to identify those library members (particularly chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (*i.e.*, the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, *e.g.*, U.S. Patent 5,010,175; Furka, *Int J Pept Prot Res* 37:487-493, 1991; and Houghton *et al.*, *Nature* 354:84-88, 1991). Other chemistries for generating chemical diversity libraries can also be used. Such

chemistries include, but are not limited to: peptides (*e.g.*, PCT Publication No. WO 91/19735), encoded peptides (*e.g.*, PCT Publication WO 93/20242), random bio-oligomers (*e.g.*, PCT Publication No. WO 92/00091), benzodiazepines (*e.g.*, U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al., PNAS USA* 90:6909-6913, 1993), vinylogous polypeptides (Hagihara *et al., J Amer Chem Soc* 114:6568, 1992), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann *et al., J Amer Chem Soc* 114:9217-9218, 1992), analogous organic syntheses of small compound libraries (Chen *et al., J Amer Chem Soc* 116:2661, 1994), oligocarbamates (Cho *et al.* 1993 *Science* 261:1303), and/or peptidyl phosphonates (Campbell *et al., J Org Chem* 59:658, 1994), nucleic acid libraries (see Sambrook *et al. Molecular Cloning, A Laboratory Manual*, Cold Springs Harbor Press, N.Y., 1989; and Ausubel *et al. Current Protocols in Molecular Biology* Green Publishing Associates and Wiley Interscience, N.Y., 1989), peptide nucleic acid libraries (see, *e.g.*, U.S. Patent 5,539,083), antibody libraries (see, *e.g.*, Vaughn *et al. Nature Biotechnology* 14:309-314, 1996; and PCT/US96/10287), carbohydrate libraries (see, *e.g.*, Liang *et al., Science* 274:1520-1522, 1996; and U.S. Patent 5,593,853), small organic molecule libraries (see, *e.g.*, benzodiazepines, Baum 1993 C&EN, Jan 18, page 33; isoprenoids, U.S. Patent 5,569,588; thiazolidionones and methathiazones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, *e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville, KY; Symphony, Rainin, Woburn, MA; 433A Applied Biosystems, Foster City, CA; 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (see, *e.g.*, ComGenex, Princeton, N.J.; Tripos, Inc., St. Louis, MO; 3D Pharmaceuticals, Exton, PA; Martek Biosciences, Columbia, MD; etc.).

C. Solid State and Soluble High Throughput Assays

In one embodiment the invention provide soluble assays using molecules such as a domain such as a ligand binding domain, an extracellular domain, a

transmembrane domain, a transmembrane domain and a cytoplasmic domain, an active site, a subunit association region, *etc.*; a domain that is covalently linked to a heterologous protein to create a chimeric molecule; a T2R bitter taste receptor protein variant; or a cell or tissue expressing a T2R bitter taste receptor protein variant, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the domain, chimeric molecule, T2R bitter taste receptor protein variant, or cell or tissue expressing a specific T2R bitter taste receptor variant is attached to a solid phase substrate.

In the high throughput assays of the invention, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (*e.g.*, 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay several different plates per day; assay screens for up to about 6,000-20,000 different compounds are possible using the integrated systems of the invention. More recently, microfluidic approaches to reagent manipulation have been developed.

The molecule of interest can be bound to the solid state component, directly or indirectly, via covalent or non-covalent linkage, *e.g.*, via a tag. The tag can be any of a variety of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest (*e.g.*, the taste transduction molecule of interest) is attached to the solid support by interaction of the tag and the tag binder.

A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, *etc.*). Antibodies to molecules with natural binders such as biotin

are also widely available and appropriate tag binders; see, SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis, MO.

Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies are commercially available and many additional antibodies are described in the literature. For example, in one common configuration, the tag is a first antibody and the tag binder is a second antibody which recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs. For example, agonists and antagonists of cell membrane receptors (*e.g.*, cell receptor-ligand interactions such as transferrin, viral receptor ligands, cytokine receptors, chemokine receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherein family, the integrin family, the selectin family, and the like; (see, *e.g.*, Pigott & Power 1993 *The Adhesion Molecule Facts Book I*). Similarly, toxins and venoms, viral epitopes, hormones (*e.g.*, opiates, steroids, etc.), intracellular receptors (*e.g.*, which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly gly sequences of between about 5 and 200 amino acids. Such flexible linkers are known to persons of skill in the art. For example, poly(ethelyne glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid substrates are commonly derivatized or functionalized by

exposing all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is well described in the literature. See, *e.g.*, Merrifield, *J Am Chem Soc* 85:2149-2154, 1963 (describing solid phase synthesis of, *e.g.*, peptides); Geysen *et al.*, *J Immun Meth* 102:259-274, 1987 (describing synthesis of solid phase components on pins); Frank & Doring, *Tetrahedron* 44:6031-6040, 1988 (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al.*, *Science* 251:767-777, 1991; Sheldon *et al.*, *Clinical Chemistry* 39:718-719, 1993; and Kozal *et al.*, *Nature Medicine* 2:753759, 1996 (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

D. Computer-based Assays

Yet another assay for compounds that modulate taste receptor protein activity involves computer assisted drug design, in which a computer system is used to generate a three-dimensional structure of a target taste receptor protein based on the structural information encoded by its amino acid sequence. The input amino acid sequence interacts directly and actively with a preestablished algorithm in a computer program to yield secondary, tertiary, and quaternary structural models of the protein. The models of the protein structure are then examined to identify regions of the structure that have the ability to bind, *e.g.*, ligands. These regions are then used to identify ligands that bind to the protein.

The three-dimensional structural model of the protein is generated by entering protein amino acid sequences of at least 10 amino acid residues or corresponding nucleic acid sequences encoding a T2R bitter taste receptor polypeptide allelic variant into the computer system. The nucleotide sequence encoding the polypeptide, or the amino acid sequence thereof, can be any of the allelic variant taste receptors described. The amino acid sequence represents the

primary sequence or subsequence of the protein, which encodes the structural information of the protein. At least 10 residues of the amino acid sequence (or a nucleotide sequence encoding 10 amino acids) are entered into the computer system from computer keyboards, computer readable substrates that include, but are not limited to, electronic storage media (*e.g.*, magnetic diskettes, tapes, cartridges, and chips), optical media (*e.g.*, CD ROM), information distributed by internet sites, and by RAM. The three-dimensional structural model of the protein is then generated by the interaction of the amino acid sequence and the computer system, using software known to those of skill in the art.

The amino acid sequence represents a primary structure that encodes the information necessary to form the secondary, tertiary and quaternary structure of the protein of interest. The software looks at certain parameters encoded by the primary sequence to generate the structural model. These parameters are referred to as "energy terms," and primarily include electrostatic potentials, hydrophobic potentials, solvent accessible surfaces, and hydrogen bonding. Secondary energy terms include van der Walls potentials. Biological molecules form the structures that minimize the energy terms in a cumulative fashion. The computer program is therefore using these terms encoded by the primary structure or amino acid sequence to create the secondary structural model. An example for G-protein cell receptors is presented in Vaidehi *et al.* (*PNAS* 99:15308-15312, 2002), which is incorporated herein by reference in its entirety.

The tertiary structure of the protein encoded by the secondary structure is then formed on the basis of the energy terms of the secondary structure. The user at this point can enter additional variables such as whether the protein is membrane bound or soluble, its location in the body, and its cellular location, *e.g.*, cytoplasmic, surface, or nuclear. These variable along with the energy terms of the secondary structure are used to form the model of the tertiary structure. In modeling the tertiary structure, the computer program matches hydrophobic faces of secondary structure with like, and hydrophilic faces of secondary structure with like.

Once the structure has been granted, potential ligand binding regions are identified by the computer system. Three-dimensional structures for potential ligands are generated by entering amino acid or nucleotide sequences or chemical

formulas of compounds, as described above. The three-dimensional structure of the potential ligand is then compared to that of the target taste receptor protein variant to identify ligands that bind to the protein. Binding affinity between the protein and ligands is determined using energy terms to determine which ligands have an enhanced probability of binding to the protein.

Example 16: Pharmaceutical Preparations and Methods of Administration

Taste modulators can be administered directly to the mammalian subject for modulation of taste, *e.g.*, modulation of bitter taste, *in vivo*. Administration is by any of the routes normally used for introducing a modulator compound into ultimate contact with the tissue to be treated, optionally the tongue or mouth. The taste modulators are administered in any suitable manner, optionally with pharmaceutically acceptable carriers. Suitable methods of administering such modulators are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions of the present invention (see, *e.g.*, *Remington's Pharmaceutical Sciences*, 17th ed. 1985).

Formulations suitable for administration include aqueous and non-aqueous solutions, isotonic sterile solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, orally. The formulations of compounds can be presented in unit-dose or multi-dose sealed containers, such as ampoules and vials. Solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. The modulators can also be administered as part of a prepared food or drug.

The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial response in the subject over time. The dose will be determined by the efficacy of the particular taste modulators employed and the condition of the subject, as well as the body weight or surface area of the area to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular compound in a particular subject.

In determining the effective amounts of the modulator to be administered, a physician may evaluate circulating plasma levels of the modulator, modulator toxicities, and the production of anti-modulator antibodies. In general, the dose equivalent of a modulator is from about 1 ng/kg to 10 mg/kg for a typical subject.

For administration, taste modulators of the present invention can be administered at a rate determined by the LD_{50} of the modulator, and the side effects of the inhibitor at various concentrations, as applied to the mass and overall health of the subject. Administration can be accomplished via single or divided doses.

This disclosure provides a comprehensive collection of SNPs in bitter taste receptor genes, including sub-sets that represent conserved, non-conserved, silent, and truncation mutations in the corresponding proteins, and individual allelic sequences for bitter taste receptor genes. The disclosure further provides methods for using the corresponding allelic variants of the taste receptor genes, alone or in various combinations, to test a subject's bitter tasting profile and to identify and analyze compounds that interact with and/or influence bitter tastes in subjects. It will be apparent that the precise details of the methods described may be varied or modified without departing from the spirit of the described invention. We claim all such modifications and variations that fall within the scope and spirit of this disclosure.

ABSTRACT

FUNCTIONAL VARIANTS OF HUMAN TASTE RECEPTOR GENES

We have identified different forms of bitter receptor genes that occur in different humans. These alleles are generated by numerous coding Single Nucleotide Polymorphisms (SNP's) that occur within the members of this gene family that exists in humans. Some SNP's cause amino acid substitutions, while others introduce chain termination codons, rendering the allele non-functional. The frequency of individual SNP alleles is often high, indicating that large numbers of individuals in the population will carry these different bitter taste receptors. Differences in these genes can have a large effect on those individuals' sense of bitter taste. This means that these individuals perceive the taste of bitter substances differently than the rest of the population. The ability to assay this allelic information is useful in the development of flavorings and flavor enhancers, as it can be used to define large groups within the population who perceive bitter tastes differently. This in turn allows the taste preferences of these groups to be addressed for the first time.

GENE	RNA		BAC CLONE	SIZE (BP)	(AA)	POSITION	SNP	GENOME	RNA	PROTEIN	ALLELE FREQUENCY	GENOTYPE FREQUENCY
T2R1	AF227129	NM_019599	AC026787	900	299	48928 - 49827 (-)						
							rs2234231	C/T	128	Pro/Leu (43)	•	
							rs41469	A/G	332	His/Arg (111)	43 : 3 (0.93 : 0.07)	A/A (0) : G/A (3): G/G (20)
							rs2234232	G/A	422	Cys/Tyr (141)	•	
							rs2234233	C/T	616	Arg/Trp (206)	•	
							rs2234234	C/T	675	Ser/Ser (225)	•	
							rs2234235	T/C	850	Leu/Leu (284)	•	
T2R3	AF227130	NM_016943	AC004979.1	951	316	17576 - 18526 (+)						
							NEW	C/T	349	Pro/Ser (117)	1	
							rs2270009	C/T	807	Gly/Gly (269)	36 : 10 (0.78 : 0.22)	C/C (13) : C/T (10) : T/T (0)
							NEW	C/T	852	Leu/Leu (284)	1	
T2R4	AF227131	NM_016944	AC004979.1	900	299	31906 - 31805 (+)						
							rs2233995	G/A	8	Arg/Gln (3)	1	
							rs2233996	G/C	9	Arg/Arg (3)	2	
							rs2233997	A/C	17	Tyr/Ser (6)	•	
							rs2233998	T/C	20	Phe/Ser (7)	31 : 7 (0.82 : 0.12)	T/T (12) : T/C (7) : C/C (0)
							rs2233999	T/A	186	Phe/Leu (62)	2 (only R seq)	
							rs2234000	C/T	221	Thr/Met (74)	•	
							rs2234001	G/C	286	Val/Leu (96)	24 : 22 (0.52 : 0.48)	G/G (6) : G/C (12) : C/C (5)
							rs2234002	G/A	512	Ser/Asn (171)	33 : 13 (0.72 : 0.28)	G/G (11) : G/A (11) : A/A (1)
							rs2234003	G/A	571	Ile/Val (191)	•	
T2R5	AF227132	NM_018980	AC004979.1	900	299	43779 - 44678 (+)						
							rs2234013	G/A	58	Gly/Ser (20)	•	
							rs2227264	G/T	77	Ser/Ile (26)	28 : 12 (0.7 : 0/3)	G/G (9) : G/T (10) : T/T (1)
							NEW	C/T	235	Arg/Cys (79)	1	
							rs2234014	C/T	338	Pro/Leu (113)	2 (only R seq)	
							NEW	G/A	363	Arg/Arg (121)	42 : 4 (0.91 : 0.09)	G/G (19) : G/A (4) : A/A (0)
							NEW	G/A	500	Cys/Tyr (167)	42 : 4 (0.91 : 0.09)	G/G (19) : G/A (4) : A/A (0)
							rs2234015	G/A	638	Arg/Gln (213)	42 : 4 (0.91 : 0.09)	G/G (19) : G/A (4) : A/A (0)
							rs2234016	G/T	881	Arg/Leu (294)	1	

GENE	RNA		BAC CLONE	SIZE (BP)	(AA)	POSITION	SNP	GENOME	RNA	PROTEIN	ALLELE FREQUENCY	GENOTYPE FREQUENCY
T2R7	AF227133	NM_023919	AC006518.17	957	318	525-1481 (-)						
							rs3759251	A/T	787	Thr/Ser (263)	*	
							NEW	C/T	788	Thr/Met (263)	43 : 3 (0.93 : 0.07)	C/C (20) : C/T (3) : T/T (0)
							rs3759252	C/A	828	Ile/Ile (276)	*	
							rs619381	G/A	912	Met/Ile (304)	*	
T2R8	AF227134	NM_023918	AC006518.17	930	309	4962 - 5891 (-)						
							NEW	A/G	496	Arg/Gly (166)	2	
							rs1548803	G/A	549	Leu/Leu (183)	40 : 6 (0.87 : 0.13)	G/G (19) : G/A (2) : A/A (2)
							NEW	T/C	829	Tyr/His (277)	1	
							rs2537817	G/A	922	Met/Val (308)	24 : 2 (0.55 : 0.45)	G/G (6) : G/A (12) : A/A (4)
T2R9	AF227135	NM_023917	AC006518.17	939	312	8048 - 8986 (-)						
							NEW	C/A	201	Phe/Leu (67)	43 : 3 (0.93 : 0.07)	C/C (21) : C/A (1) : A/A (1)
							NEW	T/A	450	Asp/Glu (150)	41 : 3 (0.93 : 0.07)	T/T (19) : T/A (3) : A/A (0)
							rs3741845	T/C	560	Val/Alu (187)	34 : 10 (0.77 : 0.23)	T/T (15) : T/C (4) : C/C (3)
							NEW	G/T	867	Leu/Phe (289)	1	
							NEW	C/A	880	Leu/Met (294)	1	
							rs3944035	C/T	910	Leu/Phe (304)	*	
							rs2159903	C/T	926	Pro/Leu (309)	*	
T2R10	AF227136	NM_023921	AC006518.17	924	307	24257 - 25180 (-)						
							NEW	A/G	120	Leu/Leu (40)	1	
							rs597468	C/T	467	Thr/Met (156)	10 : 36 (0.22 : 0.78)	C/C (0) : C/T (10) : T/T (13)
							NEW	A/C	521	Lys/Thr (174)	1	
							NEW	A/G	564	Leu/Leu (188)	2	
							NEW	G/A	627	Ser/Ser (209)	35 : 11 (0.76 : 0.24)	G/G (12) : G/A (11) : A/A (0)
T2R13	AF227137	NM_023920	AC006518.17	912	303	107298 - 108209 (-)						
							rs1015443	A/G	776	Asn/Ser (259)	40 : 6 (0.87 : 0.13)	A/A (19) : A/G (2) : G/G (2)

GENE	RNA		BAC CLONE	SIZE (BP)	(AA)	POSITION	SNP	GENOME	RNA	PROTEIN	ALLELE FREQUENCY	GENOTYPE FREQUENCY
T2R14	AF227138	NM_023922	AC006518.17	954	317	137165 - 138118 (-)						
							NEW	A/G	256	Thr/Ala (86)	2	
							rs3741843	G/A	375	Arg/Arg (125)	26 : 12 (0.68 : 0:32)	G/G (9) : G/A (8) : A/A (2)
T2R16	AF227139	NM_016945	AC004838.2	876	291	10968 - 11843 (+)						
							rs2233988	C/T	300	Thr/Thr (100)	26 : 12 (0.68 : 0.32)	C/C (9) : C/T (8) : T/T (2)
							NEW	G/A	301	Val/Met (101)	1	
							rs2692396	G/C	303	Val/Val (101)	26 : 12 (0.68 : 0.32)	G/C (9) : G/C (8) : C/C (2)
							rs2233989	T/C	460	Leu/Leu (154)	34 : 4 (0.89 : 0.11)	T/T (15) : T/C (4)N : 4 (0)
							rs846664	T/G	516	Asn/Lyn (172)	26 : 12 (0.68 : 0.32)	T/T (9) : T/G (8) : G/G (2)
							rs860170	G/A	665	Arg/His (222)	7 : 39 (0.15 : 0.85)	G/G (0) : G/A (7) : A/A (16)
							rs1204014	G/A	846	Thr/Thr (282)	26 : 12 (0.68 : 0.32)	G/G (11) : G/A (9) : A/A (3)
T2R38	AF494231	NM_176817	AC073647.9	1002	333	55139 - 56130 (+)						
							rs713598	C/G	145	Pro/Ala (49)		
							NEW	G/A	239	Arg/His (80)	1	
							rs1726866	C/T	785	Ala/Val (262)		
							NEW	C/T	820	Arg/Cys (274)	1 or 2	
							NEW	G/A	886	Val/Ile (296)		
T2R39 (PTC07)	AF494230		AC073342.3	014 (1017)	337 (333)	112 - 1117 (+)						
							NEW	C/T	578	Ser/Phe (193)	41 : 3 (0.93 : 0.07)	C/C (19) : T/C (3) : T/T (0)
							NEW	G/A	589	Lys/Glu (197)	41 : 3 (0.93 : 0.07)	A/A (19) : A/G (3) : G/G (0)
							rs4103817	G/A	874	Met/Ile (258)	*	
T2R40	AF494229		AC073342.3	972	323	38761 - 39732 (+)						
							NEW	C/A	560	Ser/Tyr (187)	38 : 8 (0.83 : 0.17)	C/C (17) : C/A (4) : A/A (2)
							NEW	G/A	817	Thr/Ala (273)	39 : 3 (0.93 : 0.07)	G/G (0) : G/A (3) : A/A (18)
T2R41 (PTC01)	AF494232		AC073264.5	924	307	27192 - 28115 (+)						
							rs1404635	G/A	189	Thr/Thr (63)	1	
							NEW	C/T	380	Pro/Leu (127)	1	
							NEW	T/A	584	Val/Asp (195)	1	

GENE	RNA	BAC CLONE	SIZE (BP)	(AA)	POSITION	SNP	GENOME	RNA	PROTEIN	ALLELE FREQUENCY	GENOTYPE FREQUENCY
T2R43	AF494237	AC018630.40	930	309	48459 - 49388 (+)						
						NEW	G/C	104	Trp/Ser (35)	30 : 4 (0.88 : 0.12)	G/G (15) : G/C (0) : C/C (2)
						NEW	G/A	270	Val/Val (90)	1 (RC11)	
						NEW	G/C	460	Arg/Gly (145)	1 (RC11)	
						NEW	T/G	510	Ser/Arg (170)	1 (RC11)	
						NEW	G/T	599	Cys/Phe (200)	1 (RC11)	
						NEW	G/A	635	Arg/His (212)	4 : 30 (0.12 : 0.88)	G/G (2) : G/A (0) : A/A (20)
						NEW	C/G	663	Thr/Thr (221)	31 : 3 (0.91 : 0.09)	C/C (15) : G/G (1) : G/G (1)
						NEW	T/G	882	Phe/Leu (294)	1 (RC11)	
						NEW	T/C	883	Trp/Arg (295)	1 (RC11)	
						NEW	A/G	889	Met/Val (297)	1 (RC11)	
T2R44	AF494228	AC018630.40	930	309	109353 - 110282 (+)						
						NEW	C/T	103	Arg/Trp (35)	38 : 6 (0.84 : 0.14)	C/C (18) : C/T (2) : T/T (2)
						NEW	T/C	423	Ala/Ala (141)	5 : 39 (0.11 : 0.89)	T/T (18) : T/C (3) : C/C (1)
						NEW	T/A	484	Met/Leu (162)	34 : 10 (0.11 : 0.23)	T/T (15) : T/A (4) : A/A (2)
						NEW	G/A	599	Cys/Tyr (200)	2	
						NEW	C/G	649	Gln/Glu (217)	1	
						NEW	C/T	680	Ala/Val (227)	8 : 38 (0.17 : 0.83)	C/C (17) : C/T (4) : T/T (2)
						NEW	G/A	718	Val/Ile (240)	7 : 39 (0.43 : 0.67)	G/G (18) : G/A (3) : A/A (2)
						NEW	A/G	744	Ser/Ser (248)	3 : 43 (0.07 : 0.93)	A/A (0) : G/A (3) : G/G (20)
						NEW	C/G	827	Pro/Arg (276)	1	
						NEW	G/A	843	Trp/* (281)	1 (RC11)	
T2R45	AF494266		900	299	?						
T2R46	AF494227	AC018630.40	900 (930)	299 (309)	78394 - 79323 (+)						
						NEW	T/G	106	Phe/Val (36)	43 : 3 (0.93 : 0.07)	T/T (20) : T/G (3) : G/G (0)
						NEW	T/A	682	Leu/Met (228)	39 : 7 (0.85 : 0.15)	T/T (18) : T/A (3) : A/A (2)
						NEW	G/A	749	Trp/* (250)	40 : 6 (0.87 : 0.13)	G/G (18) : G/A (4) : A/A (1)
						NEW	C/T	862	Gln/* (288)	1 (RC11)	
T2R47	AF494233	AC018630.40	980	319	6444 - 7403 (+)						
						rs2597925	G/A	934	Gly/Gly (314)	?	
						rs2597924	G/A	920	Arg/His (307)	?	
						rs1669405	T/G	842	Leu/Trp (281)	?	
						rs2599404	T/G	756	Phe/Leu (252)	?	

GENE	RNA		BAC CLONE	SIZE (BP)	(AA)	POSITION	SNP	GENOME	RNA	PROTEIN	ALLELE FREQUENCY	GENOTYPE FREQUENCY
T2R48	AF494234		AC018630.40	900	299	118117 - 119016 (+)						
							NEW	C/T	84	Ala/Ala (28)	38 : 8 (0.83 : 017)	C/C (17) : C/T (4) : T/T (2)
							NEW	G/A	94	Val/Ile (32)	33 : 13 (0.72 : 0.28)	G/G (12) : G/A (9) : A/A (2)
							NEW	A/C	326	Lys/Thr (109)	1	
							rs1868769	T/C	418	Leu/Leu (140	12 : 34 (0.26 : 0.74)	T/T (4) : T/C (4) : C/C (15)
							NEW	A/T	456	Arg/Ser (152)	1	
							NEW	A/G	673	Ile/Val (225)	1	
							NEW	T/C	719	Ile/Thr (240)	1	
							NEW	G/C	799	Val/Leu (267)	1	
							NEW	G/A	885	Trp/* (295)	1 (PTC12)	
							NEW	C/T	895	Arg/Cys (299)	38 : 8 (0.83 : 0.17)	C/C (17) : C/T (4) : T/T (2)
T2R49	AF494236		AC018630.40	930	309	142813 - 143742 (+)						
							NEW	A/G	156	Ala/Ala (52)	6 : 32 (0.16 : 0.84)	A/A (15) : A/G (2) : G/G (2)
							NEW	C/T	261	Ala/Ala (87)	32 : 6 (0.84 : 0.16)	C/C (15) : C/T (2) : T/T (2)
							NEW	G/A	421	Val/Ile (141)	35 : 3 (0.92 : 0.08)	G/G (16) : G/A (3) : A/A (0)
							NEW	C/A	429	His/Gln (143)	32 : 6 (0.84 : 0.16)	C/C (15) : C/A (2) : A/A (2)
							NEW	C/A	442	His/Asn (148	29 : 9 (0.76 : 0.24)	C/C (15) : C/A (5) : A/A (2)
							NEW	G/A	516	Met/Ile (172)	1	
							NEW	A/G	706	Ile/Val (236)	6 : 32 (0.16 : 0.84)	A/A (15) : A/G (2) : G/G (2)
							NEW	T/C	755	Phe/Ser (252)	32 : 6 (0.84 : 0.16)	T/T (15) : T/C (2) : C/C (2)
							NEW	G/T	764	Arg/Leu (255)	32 : 6 (0.84 : 0.16)	G/G (15) : G/T (2) : T/T (2)
							NEW	A/G	808	Ile/Val (270)	2	
T2R50	AF494235		AC018630.40	900	299	153828 - 154727 (+)						
							rs1376251	G/A	608	Tyr/Cys (203)	?	
T2R60	AY114094	NM_177437	AC092214.3	957	318	62817 - 63773 (+)						
							NEW	A/T	595	Met/Leu (199)	3 : 31 (0.09 : 0.91)	A/A (13) : A/T (3) : T/T (0)
							NEW	C/T	930	Arg/Arg (310)	24 : 22 (0.52 : 0.48)	C/C (7) : C/T (8) : T/T (8)

SEQUENCE LISTING

<110> The Government of the United States of America as represented by the Secretary of the Department of Health and Human Services
Drayna, Dennis
Kim, Un-kyung

<120> FUNCTIONAL VARIANTS OF HUMAN TASTE RECEPTOR GENES

<130> 4239-66168

<160> 46

<170> PatentIn version 3.2

<210> 1
<211> 900
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(900)

<400> 1

```
atg cta gag tct cac ctc att atc tat ttt ctt ctt gca gtg ata caa      48
Met Leu Glu Ser His Leu Ile Ile Tyr Phe Leu Leu Ala Val Ile Gln
1               5                  10                  15

ttt ctt ctt ggg att ttc aca aat ggc atc att gtg gtg gtg aat ggc      96
Phe Leu Leu Gly Ile Phe Thr Asn Gly Ile Ile Val Val Val Asn Gly
            20                  25                  30

att gac ttg atc aag cac aga aaa atg gct ccg ctg gat ctc ctt ctt     144
Ile Asp Leu Ile Lys His Arg Lys Met Ala Pro Leu Asp Leu Leu Leu
        35                  40                  45

tct tgt ctg gca gtt tct aga att ttt ctg cag ttg ttc atc ttc tac     192
Ser Cys Leu Ala Val Ser Arg Ile Phe Leu Gln Leu Phe Ile Phe Tyr
    50                  55                  60

gtt aat gtg att gtt atc ttc ttc ata gaa ttc atc atg tgt tct gcg     240
Val Asn Val Ile Val Ile Phe Phe Ile Glu Phe Ile Met Cys Ser Ala
65                  70                  75                  80

aat tgt gca att ctc tta ttt ata aat gaa ttg gaa ctt tgg ctt gcc     288
Asn Cys Ala Ile Leu Leu Phe Ile Asn Glu Leu Glu Leu Trp Leu Ala
                85                  90                  95

aca tgg ctc ggc gtt ttc tat tgt gcc aag gtt gcc agc gtc crt cac     336
Thr Trp Leu Gly Val Phe Tyr Cys Ala Lys Val Ala Ser Val Xaa His
            100                 105                 110

cca ctc ttc atc tgg ttg aag atg agg ata tcc aag ctg gtc cca tgg     384
Pro Leu Phe Ile Trp Leu Lys Met Arg Ile Ser Lys Leu Val Pro Trp
        115                 120                 125

atg atc ctg ggg tct ctg cta tat gta tct atg att tgt gtt ttc cat     432
```

```
Met Ile Leu Gly Ser Leu Leu Tyr Val Ser Met Ile Cys Val Phe His
    130                 135                 140

agc aaa tat gca ggg ttt atg gtc cca tac ttc cta agg aaa ttt ttc      480
Ser Lys Tyr Ala Gly Phe Met Val Pro Tyr Phe Leu Arg Lys Phe Phe
145                 150                 155                 160

tcc caa aat gcc aca att caa aaa gaa gat aca ctg gct ata cag att      528
Ser Gln Asn Ala Thr Ile Gln Lys Glu Asp Thr Leu Ala Ile Gln Ile
                165                 170                 175

ttc tct ttt gtt gct gag ttc tca gtg cca ttg ctt atc ttc ctt ttt      576
Phe Ser Phe Val Ala Glu Phe Ser Val Pro Leu Leu Ile Phe Leu Phe
            180                 185                 190

gct gtt ttg ctc ttg att ttc tct ctg ggg agg cac acc cgg caa atg      624
Ala Val Leu Leu Leu Ile Phe Ser Leu Gly Arg His Thr Arg Gln Met
        195                 200                 205

aga aac aca gtg gcc ggc agc agg gtt cct ggc agg ggt gca ccc atc      672
Arg Asn Thr Val Ala Gly Ser Arg Val Pro Gly Arg Gly Ala Pro Ile
    210                 215                 220

agc gcg ttg ctg tct atc ctg tcc ttc ctg atc ctc tac ttc tcc cac      720
Ser Ala Leu Leu Ser Ile Leu Ser Phe Leu Ile Leu Tyr Phe Ser His
225                 230                 235                 240

tgc atg ata aaa gtt ttt ctc tct tct cta aag ttt cac atc aga agg      768
Cys Met Ile Lys Val Phe Leu Ser Ser Leu Lys Phe His Ile Arg Arg
                245                 250                 255

ttc atc ttt ctg ttc ttc atc ctt gtg att ggt ata tac cct tct gga      816
Phe Ile Phe Leu Phe Phe Ile Leu Val Ile Gly Ile Tyr Pro Ser Gly
            260                 265                 270

cac tct ctc atc tta att tta gga aat cct aaa ttg aaa caa aat gca      864
His Ser Leu Ile Leu Ile Leu Gly Asn Pro Lys Leu Lys Gln Asn Ala
        275                 280                 285

aaa aag ttc ctc ctc cac agt aag tgc tgt cag tga                      900
Lys Lys Phe Leu Leu His Ser Lys Cys Cys Gln
    290                 295
```

<210> 2
<211> 299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (111)..(111)
<223> The 'Xaa' at location 111 stands for Arg, or His.

<400> 2

```
Met Leu Glu Ser His Leu Ile Ile Tyr Phe Leu Leu Ala Val Ile Gln
1               5                   10                  15
```

Phe Leu Leu Gly Ile Phe Thr Asn Gly Ile Ile Val Val Val Asn Gly
20 25 30

Ile Asp Leu Ile Lys His Arg Lys Met Ala Pro Leu Asp Leu Leu Leu
35 40 45

Ser Cys Leu Ala Val Ser Arg Ile Phe Leu Gln Leu Phe Ile Phe Tyr
50 55 60

Val Asn Val Ile Val Ile Phe Phe Ile Glu Phe Ile Met Cys Ser Ala
65 70 75 80

Asn Cys Ala Ile Leu Leu Phe Ile Asn Glu Leu Glu Leu Trp Leu Ala
85 90 95

Thr Trp Leu Gly Val Phe Tyr Cys Ala Lys Val Ala Ser Val Xaa His
100 105 110

Pro Leu Phe Ile Trp Leu Lys Met Arg Ile Ser Lys Leu Val Pro Trp
115 120 125

Met Ile Leu Gly Ser Leu Leu Tyr Val Ser Met Ile Cys Val Phe His
130 135 140

Ser Lys Tyr Ala Gly Phe Met Val Pro Tyr Phe Leu Arg Lys Phe Phe
145 150 155 160

Ser Gln Asn Ala Thr Ile Gln Lys Glu Asp Thr Leu Ala Ile Gln Ile
165 170 175

Phe Ser Phe Val Ala Glu Phe Ser Val Pro Leu Leu Ile Phe Leu Phe
180 185 190

Ala Val Leu Leu Leu Ile Phe Ser Leu Gly Arg His Thr Arg Gln Met
195 200 205

Arg Asn Thr Val Ala Gly Ser Arg Val Pro Gly Arg Gly Ala Pro Ile
210 215 220

Ser Ala Leu Leu Ser Ile Leu Ser Phe Leu Ile Leu Tyr Phe Ser His
225 230 235 240

Cys Met Ile Lys Val Phe Leu Ser Ser Leu Lys Phe His Ile Arg Arg
245 250 255

```
Phe Ile Phe Leu Phe Phe Ile Leu Val Ile Gly Ile Tyr Pro Ser Gly
            260                 265                 270

His Ser Leu Ile Leu Ile Leu Gly Asn Pro Lys Leu Lys Gln Asn Ala
        275                 280                 285

Lys Lys Phe Leu Leu His Ser Lys Cys Cys Gln
    290                 295
```

<210> 3
<211> 951
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(951)

<400> 3

```
atg atg gga ctc acc gag ggg gtg ttc ctg att ctg tct ggc act cag      48
Met Met Gly Leu Thr Glu Gly Val Phe Leu Ile Leu Ser Gly Thr Gln
1               5                   10                  15

ttc aca ctg gga att ctg gtc aat tgt ttc att gag ttg gtc aat ggt      96
Phe Thr Leu Gly Ile Leu Val Asn Cys Phe Ile Glu Leu Val Asn Gly
            20                  25                  30

agc agc tgg ttc aag acc aag aga atg tct ttg tct gac ttc atc atc     144
Ser Ser Trp Phe Lys Thr Lys Arg Met Ser Leu Ser Asp Phe Ile Ile
        35                  40                  45

acc acc ctg gca ctc ttg agg atc att ctg ctg tgt att atc ttg act     192
Thr Thr Leu Ala Leu Leu Arg Ile Ile Leu Leu Cys Ile Ile Leu Thr
    50                  55                  60

gat agt ttt tta ata gaa ttc tct ccc aac aca cat gat tca ggg ata     240
Asp Ser Phe Leu Ile Glu Phe Ser Pro Asn Thr His Asp Ser Gly Ile
65                  70                  75                  80

ata atg caa att att gat gtt tcc tgg aca ttt aca aac cat ctg agc     288
Ile Met Gln Ile Ile Asp Val Ser Trp Thr Phe Thr Asn His Leu Ser
                85                  90                  95

att tgg ctt gcc acc tgt ctt ggt gtc ctc tac tgc ctg aaa atc gcc     336
Ile Trp Leu Ala Thr Cys Leu Gly Val Leu Tyr Cys Leu Lys Ile Ala
            100                 105                 110

agt ttc tct cac ycc aca ttc ctc tgg ctc aag tgg aga gtt tct agg     384
Ser Phe Ser His Xaa Thr Phe Leu Trp Leu Lys Trp Arg Val Ser Arg
        115                 120                 125

gtg atg gta tgg atg ctg ttg ggt gca ctg ctc tta tcc tgt ggt agt     432
Val Met Val Trp Met Leu Leu Gly Ala Leu Leu Leu Ser Cys Gly Ser
    130                 135                 140

acc gca tct ctg atc aat gag ttt aag ctc tat tct gtc ttt agg gga     480
Thr Ala Ser Leu Ile Asn Glu Phe Lys Leu Tyr Ser Val Phe Arg Gly
```

```
145                 150                 155                 160

att gag gcc acc agg aat gtg act gaa cac ttc aga aag aag agg agt      528
Ile Glu Ala Thr Arg Asn Val Thr Glu His Phe Arg Lys Lys Arg Ser
                165                 170                 175

gag tat tat ctg atc cat gtt ctt ggg act ctg tgg tac ctg cct ccc      576
Glu Tyr Tyr Leu Ile His Val Leu Gly Thr Leu Trp Tyr Leu Pro Pro
            180                 185                 190

tta att gtg tcc ctg gcc tcc tac tct ttg ctc atc ttc tcc ctg ggg      624
Leu Ile Val Ser Leu Ala Ser Tyr Ser Leu Leu Ile Phe Ser Leu Gly
        195                 200                 205

agg cac aca cgg cag atg ctg caa aat ggg aca agc tcc aga gat cca      672
Arg His Thr Arg Gln Met Leu Gln Asn Gly Thr Ser Ser Arg Asp Pro
    210                 215                 220

acc act gag gcc cac aag agg gcc atc aga atc atc ctt tcc ttc ttc      720
Thr Thr Glu Ala His Lys Arg Ala Ile Arg Ile Ile Leu Ser Phe Phe
225                 230                 235                 240

ttt ctc ttc tta ctt tac ttt ctt gct ttc tta att gca tca ttt ggt      768
Phe Leu Phe Leu Leu Tyr Phe Leu Ala Phe Leu Ile Ala Ser Phe Gly
                245                 250                 255

aat ttc cta cca aaa acc aag atg gct aag atg att ggy gaa gta atg      816
Asn Phe Leu Pro Lys Thr Lys Met Ala Lys Met Ile Xaa Glu Val Met
            260                 265                 270

aca atg ttt tat cct gct ggc cac tca ttt att cty att ctg ggg aac      864
Thr Met Phe Tyr Pro Ala Gly His Ser Phe Ile Xaa Ile Leu Gly Asn
        275                 280                 285

agt aag ctg aag cag aca ttt gta gtg atg ctc cgg tgt gag tct ggt      912
Ser Lys Leu Lys Gln Thr Phe Val Val Met Leu Arg Cys Glu Ser Gly
    290                 295                 300

cat ctg aag cct gga tcc aag gga ccc att ttc tct tag                  951
His Leu Lys Pro Gly Ser Lys Gly Pro Ile Phe Ser
305                 310                 315
```

<210> 4
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (117)..(117)
<223> The 'Xaa' at location 117 stands for Pro, or Ser.

<220>
<221> misc_feature
<222> (269)..(269)
<223> The 'Xaa' at location 269 stands for Gly.

<220>
<221> misc_feature
<222> (284)..(284)

<223> The 'Xaa' at location 284 stands for Leu.

<400> 4

Met Met Gly Leu Thr Glu Gly Val Phe Leu Ile Leu Ser Gly Thr Gln
1 5 10 15

Phe Thr Leu Gly Ile Leu Val Asn Cys Phe Ile Glu Leu Val Asn Gly
20 25 30

Ser Ser Trp Phe Lys Thr Lys Arg Met Ser Leu Ser Asp Phe Ile Ile
35 40 45

Thr Thr Leu Ala Leu Leu Arg Ile Ile Leu Leu Cys Ile Ile Leu Thr
50 55 60

Asp Ser Phe Leu Ile Glu Phe Ser Pro Asn Thr His Asp Ser Gly Ile
65 70 75 80

Ile Met Gln Ile Ile Asp Val Ser Trp Thr Phe Thr Asn His Leu Ser
85 90 95

Ile Trp Leu Ala Thr Cys Leu Gly Val Leu Tyr Cys Leu Lys Ile Ala
100 105 110

Ser Phe Ser His Xaa Thr Phe Leu Trp Leu Lys Trp Arg Val Ser Arg
115 120 125

Val Met Val Trp Met Leu Leu Gly Ala Leu Leu Leu Ser Cys Gly Ser
130 135 140

Thr Ala Ser Leu Ile Asn Glu Phe Lys Leu Tyr Ser Val Phe Arg Gly
145 150 155 160

Ile Glu Ala Thr Arg Asn Val Thr Glu His Phe Arg Lys Lys Arg Ser
165 170 175

Glu Tyr Tyr Leu Ile His Val Leu Gly Thr Leu Trp Tyr Leu Pro Pro
180 185 190

Leu Ile Val Ser Leu Ala Ser Tyr Ser Leu Leu Ile Phe Ser Leu Gly
195 200 205

Arg His Thr Arg Gln Met Leu Gln Asn Gly Thr Ser Ser Arg Asp Pro
210 215 220

Thr Thr Glu Ala His Lys Arg Ala Ile Arg Ile Ile Leu Ser Phe Phe

```
225                 230                 235                 240

Phe Leu Phe Leu Leu Tyr Phe Leu Ala Phe Leu Ile Ala Ser Phe Gly
                245                 250                 255

Asn Phe Leu Pro Lys Thr Lys Met Ala Lys Met Ile Xaa Glu Val Met
            260                 265                 270

Thr Met Phe Tyr Pro Ala Gly His Ser Phe Ile Xaa Ile Leu Gly Asn
        275                 280                 285

Ser Lys Leu Lys Gln Thr Phe Val Val Met Leu Arg Cys Glu Ser Gly
    290                 295                 300

His Leu Lys Pro Gly Ser Lys Gly Pro Ile Phe Ser
305                 310                 315
```

<210> 5
<211> 900
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(900)

<400> 5

```
atg ctt crs tta ttc tat tyc tct gct att att gcc tca gtt att tta      48
Met Leu Xaa Leu Phe Tyr Xaa Ser Ala Ile Ile Ala Ser Val Ile Leu
1               5                   10                  15

aat ttt gta gga atc att atg aat ctg ttt att aca gtg gtc aat tgc      96
Asn Phe Val Gly Ile Ile Met Asn Leu Phe Ile Thr Val Val Asn Cys
            20                  25                  30

aaa act tgg gtc aaa agc cat aga atc tcc tct tct gat agg att ctg     144
Lys Thr Trp Val Lys Ser His Arg Ile Ser Ser Ser Asp Arg Ile Leu
        35                  40                  45

ttc agc ctg ggc atc acc agg ttt ctt atg ctg gga cta ttw ctg gtg     192
Phe Ser Leu Gly Ile Thr Arg Phe Leu Met Leu Gly Leu Xaa Leu Val
    50                  55                  60

aac acc atc tac ttc gtc tct tca aat acg gaa agg tca gtc tac ctg     240
Asn Thr Ile Tyr Phe Val Ser Ser Asn Thr Glu Arg Ser Val Tyr Leu
65                  70                  75                  80

tct gct ttt ttt gtg ttg tgt ttc atg ttt ttg gac tcg agc agt stc     288
Ser Ala Phe Phe Val Leu Cys Phe Met Phe Leu Asp Ser Ser Ser Xaa
                85                  90                  95

tgg ttt gtg acc ttg ctc aat atc ttg tac tgt gtg aag att act aac     336
Trp Phe Val Thr Leu Leu Asn Ile Leu Tyr Cys Val Lys Ile Thr Asn
            100                 105                 110
```

```
ttc caa cac tca gtg ttt ctc ctg ctg aag cgg aat atc tcc cca aag      384
Phe Gln His Ser Val Phe Leu Leu Leu Lys Arg Asn Ile Ser Pro Lys
        115                 120                 125

atc ccc agg ctg ctg ctg gcc tgt gtg ctg att tct gct ttc acc act      432
Ile Pro Arg Leu Leu Leu Ala Cys Val Leu Ile Ser Ala Phe Thr Thr
    130                 135                 140

tgc ctg tac atc acg ctt agc cag gca tca cct ttt cct gaa ctt gtg      480
Cys Leu Tyr Ile Thr Leu Ser Gln Ala Ser Pro Phe Pro Glu Leu Val
145                 150                 155                 160

act acg aga aat aac aca tca ttt aat atc art gag ggc atc ttg tct      528
Thr Thr Arg Asn Asn Thr Ser Phe Asn Ile Xaa Glu Gly Ile Leu Ser
                165                 170                 175

tta gtg gtt tct ttg gtc ttg agc tca tct ctc cag ttc atc att aat      576
Leu Val Val Ser Leu Val Leu Ser Ser Ser Leu Gln Phe Ile Ile Asn
            180                 185                 190

gtg act tct gct tcc ttg cta ata cac tcc ttg agg aga cat ata cag      624
Val Thr Ser Ala Ser Leu Leu Ile His Ser Leu Arg Arg His Ile Gln
        195                 200                 205

aag atg cag aaa aat gcc act ggt ttc tgg aat ccc cag acg gaa gct      672
Lys Met Gln Lys Asn Ala Thr Gly Phe Trp Asn Pro Gln Thr Glu Ala
    210                 215                 220

cat gta ggt gct atg aag ctg atg gtc tat ttc ctc atc ctc tac att      720
His Val Gly Ala Met Lys Leu Met Val Tyr Phe Leu Ile Leu Tyr Ile
225                 230                 235                 240

cca tat tca gtt gct acc ctg gtc cag tat ctc ccc ttt tat gca ggg      768
Pro Tyr Ser Val Ala Thr Leu Val Gln Tyr Leu Pro Phe Tyr Ala Gly
                245                 250                 255

atg gat atg ggg acc aaa tcc att tgt ctg att ttt gcc acc ctt tac      816
Met Asp Met Gly Thr Lys Ser Ile Cys Leu Ile Phe Ala Thr Leu Tyr
            260                 265                 270

tct cca gga cat tct gtt ctc att att atc aca cat cct aaa ctg aaa      864
Ser Pro Gly His Ser Val Leu Ile Ile Ile Thr His Pro Lys Leu Lys
        275                 280                 285

aca aca gca aag aag att ctt tgt ttc aaa aaa tag                      900
Thr Thr Ala Lys Lys Ile Leu Cys Phe Lys Lys
    290                 295
```

<210> 6
<211> 299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (3)..(3)
<223> The 'Xaa' at location 3 stands for Arg, Gln, or His.

<220>

<221> misc_feature
<222> (7)..(7)
<223> The 'Xaa' at location 7 stands for Ser, or Phe.

<220>
<221> misc_feature
<222> (62)..(62)
<223> The 'Xaa' at location 62 stands for Leu, or Phe.

<220>
<221> misc_feature
<222> (96)..(96)
<223> The 'Xaa' at location 96 stands for Val, or Leu.

<220>
<221> misc_feature
<222> (171)..(171)
<223> The 'Xaa' at location 171 stands for Ser, or Asn.

<400> 6

```
Met Leu Xaa Leu Phe Tyr Xaa Ser Ala Ile Ile Ala Ser Val Ile Leu
1               5                  10                  15

Asn Phe Val Gly Ile Ile Met Asn Leu Phe Ile Thr Val Val Asn Cys
            20                  25                  30

Lys Thr Trp Val Lys Ser His Arg Ile Ser Ser Ser Asp Arg Ile Leu
        35                  40                  45

Phe Ser Leu Gly Ile Thr Arg Phe Leu Met Leu Gly Leu Xaa Leu Val
    50                  55                  60

Asn Thr Ile Tyr Phe Val Ser Ser Asn Thr Glu Arg Ser Val Tyr Leu
65                  70                  75                  80

Ser Ala Phe Phe Val Leu Cys Phe Met Phe Leu Asp Ser Ser Ser Xaa
                85                  90                  95

Trp Phe Val Thr Leu Leu Asn Ile Leu Tyr Cys Val Lys Ile Thr Asn
            100                 105                 110

Phe Gln His Ser Val Phe Leu Leu Leu Lys Arg Asn Ile Ser Pro Lys
        115                 120                 125

Ile Pro Arg Leu Leu Leu Ala Cys Val Leu Ile Ser Ala Phe Thr Thr
    130                 135                 140

Cys Leu Tyr Ile Thr Leu Ser Gln Ala Ser Pro Phe Pro Glu Leu Val
145                 150                 155                 160
```

```
Thr Thr Arg Asn Asn Thr Ser Phe Asn Ile Xaa Glu Gly Ile Leu Ser
                165                 170                 175

Leu Val Val Ser Leu Val Leu Ser Ser Ser Leu Gln Phe Ile Ile Asn
            180                 185                 190

Val Thr Ser Ala Ser Leu Leu Ile His Ser Leu Arg Arg His Ile Gln
        195                 200                 205

Lys Met Gln Lys Asn Ala Thr Gly Phe Trp Asn Pro Gln Thr Glu Ala
    210                 215                 220

His Val Gly Ala Met Lys Leu Met Val Tyr Phe Leu Ile Leu Tyr Ile
225                 230                 235                 240

Pro Tyr Ser Val Ala Thr Leu Val Gln Tyr Leu Pro Phe Tyr Ala Gly
                245                 250                 255

Met Asp Met Gly Thr Lys Ser Ile Cys Leu Ile Phe Ala Thr Leu Tyr
            260                 265                 270

Ser Pro Gly His Ser Val Leu Ile Ile Ile Thr His Pro Lys Leu Lys
        275                 280                 285

Thr Thr Ala Lys Lys Ile Leu Cys Phe Lys Lys
    290                 295
```

<210> 7
<211> 900
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(900)

<400> 7

```
atg ctg agc gct ggc cta gga ctg ctg atg ctg gtg gca gtg gtt gaa       48
Met Leu Ser Ala Gly Leu Gly Leu Leu Met Leu Val Ala Val Val Glu
1               5                   10                  15

ttt ctc atc ggt tta att gga aat gga akc ctg gtg gtc tgg agt ttt       96

Phe Leu Ile Gly Leu Ile Gly Asn Gly Xaa Leu Val Val Trp Ser Phe
            20                  25                  30

aga gaa tgg atc aga aaa ttc aac tgg tcc tca tat aac ctc att atc      144
Arg Glu Trp Ile Arg Lys Phe Asn Trp Ser Ser Tyr Asn Leu Ile Ile
        35                  40                  45
```

```
ctg ggc ctg gct ggc tgc cga ttt ctc ctg cag tgg ctg atc att ttg      192
Leu Gly Leu Ala Gly Cys Arg Phe Leu Leu Gln Trp Leu Ile Ile Leu
    50                  55                  60

gac tta agc ttg ttt cca ctt ttc cag agc agc cgt tgg ctt ygc tat      240
Asp Leu Ser Leu Phe Pro Leu Phe Gln Ser Ser Arg Trp Leu Xaa Tyr
65                  70                  75                  80

ctt agt atc ttc tgg gtc ctg gta agc cag gcc agc tta tgg ttt gcc      288
Leu Ser Ile Phe Trp Val Leu Val Ser Gln Ala Ser Leu Trp Phe Ala
                85                  90                  95

acc ttc ctc agt gtc ttc tat tgc aag aag atc acg acc ttc gat cgc      336
Thr Phe Leu Ser Val Phe Tyr Cys Lys Lys Ile Thr Thr Phe Asp Arg
            100                 105                 110

cyg gcc tac ttg tgg ctg aag cag agr gcc tat aac ctg agt ctc tgg      384
Xaa Ala Tyr Leu Trp Leu Lys Gln Arg Ala Tyr Asn Leu Ser Leu Trp
        115                 120                 125

tgc ctt ctg ggc tac ttt ata atc aat ttg tta ctt aca gtc caa att      432
Cys Leu Leu Gly Tyr Phe Ile Ile Asn Leu Leu Leu Thr Val Gln Ile
    130                 135                 140

ggc tta aca ttc tat cat cct ccc caa gga aac agc agc att cgg tat      480

Gly Leu Thr Phe Tyr His Pro Pro Gln Gly Asn Ser Ser Ile Arg Tyr
145                 150                 155                 160

ccc ttt gaa agc tgg cag trc ctg tat gca ttt cag ctc aat tca gga      528
Pro Phe Glu Ser Trp Gln Xaa Leu Tyr Ala Phe Gln Leu Asn Ser Gly
                165                 170                 175

agt tat ttg cct tta gtg gtg ttt ctt gtt tcc tct ggg atg ctg att      576
Ser Tyr Leu Pro Leu Val Val Phe Leu Val Ser Ser Gly Met Leu Ile
            180                 185                 190

gtc tct ttg tat aca cac cac aag aag atg aag gtc cat tca gct ggt      624
Val Ser Leu Tyr Thr His His Lys Lys Met Lys Val His Ser Ala Gly
        195                 200                 205

agg agg gat gtc crg gcc aag gct cac atc act gcg ctg aag tcc ttg      672
Arg Arg Asp Val Xaa Ala Lys Ala His Ile Thr Ala Leu Lys Ser Leu
    210                 215                 220

ggc tgc ttc ctc tta ctt cac ctg gtt tat atc atg gcc agc ccc ttc      720
Gly Cys Phe Leu Leu Leu His Leu Val Tyr Ile Met Ala Ser Pro Phe
225                 230                 235                 240

tcc atc acc tcc aag act tat cct cct gat ctc acc agt gtc ttc atc      768
Ser Ile Thr Ser Lys Thr Tyr Pro Pro Asp Leu Thr Ser Val Phe Ile
                245                 250                 255

tgg gag aca ctc atg gca gcc tat cct tct ctt cat tct ctc ata ttg      816
Trp Glu Thr Leu Met Ala Ala Tyr Pro Ser Leu His Ser Leu Ile Leu
            260                 265                 270

atc atg ggg att cct agg gtg aag cag act tgt cag aag atc ctg tgg      864

Ile Met Gly Ile Pro Arg Val Lys Gln Thr Cys Gln Lys Ile Leu Trp
        275                 280                 285
```

```
aag aca gtg tgt gct ckg aga tgc tgg ggc cca tga                          900
Lys Thr Val Cys Ala Xaa Arg Cys Trp Gly Pro
    290                 295
```

<210> 8
<211> 299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (26)..(26)
<223> The 'Xaa' at location 26 stands for Ser, or Ile.

<220>
<221> misc_feature
<222> (79)..(79)
<223> The 'Xaa' at location 79 stands for Arg, or Cys.

<220>
<221> misc_feature
<222> (113)..(113)
<223> The 'Xaa' at location 113 stands for Pro, or Leu.

<220>
<221> misc_feature
<222> (167)..(167)
<223> The 'Xaa' at location 167 stands for Cys, or Tyr.

<220>
<221> misc_feature
<222> (213)..(213)
<223> The 'Xaa' at location 213 stands for Arg, or Gln.

<220>
<221> misc_feature
<222> (294)..(294)
<223> The 'Xaa' at location 294 stands for Arg, or Leu.

<400> 8

```
Met Leu Ser Ala Gly Leu Gly Leu Leu Met Leu Val Ala Val Val Glu
1               5                  10                  15

Phe Leu Ile Gly Leu Ile Gly Asn Gly Xaa Leu Val Val Trp Ser Phe
            20                  25                  30

Arg Glu Trp Ile Arg Lys Phe Asn Trp Ser Ser Tyr Asn Leu Ile Ile
        35                  40                  45

Leu Gly Leu Ala Gly Cys Arg Phe Leu Leu Gln Trp Leu Ile Ile Leu
    50                  55                  60

Asp Leu Ser Leu Phe Pro Leu Phe Gln Ser Ser Arg Trp Leu Xaa Tyr
65                  70                  75                  80
```

Leu Ser Ile Phe Trp Val Leu Val Ser Gln Ala Ser Leu Trp Phe Ala
85 90 95

Thr Phe Leu Ser Val Phe Tyr Cys Lys Lys Ile Thr Thr Phe Asp Arg
100 105 110

Xaa Ala Tyr Leu Trp Leu Lys Gln Arg Ala Tyr Asn Leu Ser Leu Trp
115 120 125

Cys Leu Leu Gly Tyr Phe Ile Ile Asn Leu Leu Leu Thr Val Gln Ile
130 135 140

Gly Leu Thr Phe Tyr His Pro Pro Gln Gly Asn Ser Ser Ile Arg Tyr
145 150 155 160

Pro Phe Glu Ser Trp Gln Xaa Leu Tyr Ala Phe Gln Leu Asn Ser Gly
165 170 175

Ser Tyr Leu Pro Leu Val Val Phe Leu Val Ser Ser Gly Met Leu Ile
180 185 190

Val Ser Leu Tyr Thr His His Lys Lys Met Lys Val His Ser Ala Gly
195 200 205

Arg Arg Asp Val Xaa Ala Lys Ala His Ile Thr Ala Leu Lys Ser Leu
210 215 220

Gly Cys Phe Leu Leu Leu His Leu Val Tyr Ile Met Ala Ser Pro Phe
225 230 235 240

Ser Ile Thr Ser Lys Thr Tyr Pro Pro Asp Leu Thr Ser Val Phe Ile
245 250 255

Trp Glu Thr Leu Met Ala Ala Tyr Pro Ser Leu His Ser Leu Ile Leu
260 265 270

Ile Met Gly Ile Pro Arg Val Lys Gln Thr Cys Gln Lys Ile Leu Trp
275 280 285

Lys Thr Val Cys Ala Xaa Arg Cys Trp Gly Pro
290 295

<210> 9
<211> 957
<212> DNA

<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(957)

<400> 9

```
atg gca gat aaa gtg cag act act tta ttg ttc tta gca gtt gga gag      48
Met Ala Asp Lys Val Gln Thr Thr Leu Leu Phe Leu Ala Val Gly Glu
1               5                   10                  15

ttt tca gtg ggg atc tta ggg aat gca ttc att gga ttg gta aac tgc      96
Phe Ser Val Gly Ile Leu Gly Asn Ala Phe Ile Gly Leu Val Asn Cys
            20                  25                  30

atg gac tgg gtc aag aag agg aaa att gcc tcc att gat tta atc ctc     144
Met Asp Trp Val Lys Lys Arg Lys Ile Ala Ser Ile Asp Leu Ile Leu
        35                  40                  45

aca agt ctg gcc ata tcc aga att tgt cta ttg tgc gta ata cta tta     192
Thr Ser Leu Ala Ile Ser Arg Ile Cys Leu Leu Cys Val Ile Leu Leu
    50                  55                  60

gat tgt ttt ata ttg gtg cta tat cca gat gtc tat gcc act ggt aaa     240
Asp Cys Phe Ile Leu Val Leu Tyr Pro Asp Val Tyr Ala Thr Gly Lys
65                  70                  75                  80

gaa atg aga atc att gac ttc ttc tgg aca cta acc aat cat tta agt     288
Glu Met Arg Ile Ile Asp Phe Phe Trp Thr Leu Thr Asn His Leu Ser
                85                  90                  95

atc tgg ttt gca acc tgc ctc agc att tac tat ttc ttc aag ata ggt     336
Ile Trp Phe Ala Thr Cys Leu Ser Ile Tyr Tyr Phe Phe Lys Ile Gly
            100                 105                 110

aat ttc ttt cac cca ctt ttc ctc tgg atg aag tgg aga att gac agg     384
Asn Phe Phe His Pro Leu Phe Leu Trp Met Lys Trp Arg Ile Asp Arg
        115                 120                 125

gtg att tcc tgg att cta ctg ggg tgc gtg gtt ctc tct gtg ttt att     432
Val Ile Ser Trp Ile Leu Leu Gly Cys Val Val Leu Ser Val Phe Ile
    130                 135                 140

agc ctt cca gcc act gag aat ttg aac gct gat ttc agg ttt tgt gtg     480
Ser Leu Pro Ala Thr Glu Asn Leu Asn Ala Asp Phe Arg Phe Cys Val
145                 150                 155                 160

aag gca aag agg aaa aca aac tta act tgg agt tgc aga gta aat aaa     528
Lys Ala Lys Arg Lys Thr Asn Leu Thr Trp Ser Cys Arg Val Asn Lys
                165                 170                 175

act caa cat gct tct acc aag tta ttt ctc aac ctg gca acg ctg ctc     576
Thr Gln His Ala Ser Thr Lys Leu Phe Leu Asn Leu Ala Thr Leu Leu
            180                 185                 190

ccc ttt tgt gtg tgc cta atg tcc ttt ttc ctc ttg atc ctc tcc ctg     624
Pro Phe Cys Val Cys Leu Met Ser Phe Phe Leu Leu Ile Leu Ser Leu
        195                 200                 205
```

```
cgg aga cat atc agg cga atg cag ctc agt gcc aca ggg tgc aga gac      672
Arg Arg His Ile Arg Arg Met Gln Leu Ser Ala Thr Gly Cys Arg Asp
    210                 215                 220

ccc agc aca gaa gcc cat gtg aga gcc ctg aaa gct gtc att tcc ttc      720
Pro Ser Thr Glu Ala His Val Arg Ala Leu Lys Ala Val Ile Ser Phe
225                 230                 235                 240

ctt ctc ctc ttt att gcc tac tat ttg tcc ttt ctc att gcc acc tcc      768
Leu Leu Leu Phe Ile Ala Tyr Tyr Leu Ser Phe Leu Ile Ala Thr Ser
                245                 250                 255

agc tac ttt atg cca gag ayg gaa tta gct gtg att ttt ggt gag tcc      816
Ser Tyr Phe Met Pro Glu Xaa Glu Leu Ala Val Ile Phe Gly Glu Ser
            260                 265                 270

ata gct cta atc tac ccc tca agt cat tca ttt atc cta ata ctg ggg      864
Ile Ala Leu Ile Tyr Pro Ser Ser His Ser Phe Ile Leu Ile Leu Gly
        275                 280                 285

aac aat aaa tta aga cat gca tct cta aag gtg att tgg aaa gta atg      912
Asn Asn Lys Leu Arg His Ala Ser Leu Lys Val Ile Trp Lys Val Met
    290                 295                 300

tct att cta aaa gga aga aaa ttc caa caa cat aaa caa atc tga          957
Ser Ile Leu Lys Gly Arg Lys Phe Gln Gln His Lys Gln Ile
305                 310                 315
```

<210> 10
<211> 318
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (263)..(263)
<223> The 'Xaa' at location 263 stands for Thr, or Met.

<400> 10

```
Met Ala Asp Lys Val Gln Thr Thr Leu Leu Phe Leu Ala Val Gly Glu
1               5                   10                  15

Phe Ser Val Gly Ile Leu Gly Asn Ala Phe Ile Gly Leu Val Asn Cys
            20                  25                  30

Met Asp Trp Val Lys Lys Arg Lys Ile Ala Ser Ile Asp Leu Ile Leu
        35                  40                  45

Thr Ser Leu Ala Ile Ser Arg Ile Cys Leu Leu Cys Val Ile Leu Leu
    50                  55                  60

Asp Cys Phe Ile Leu Val Leu Tyr Pro Asp Val Tyr Ala Thr Gly Lys
65                  70                  75                  80
```

Glu Met Arg Ile Ile Asp Phe Phe Trp Thr Leu Thr Asn His Leu Ser
85 90 95

Ile Trp Phe Ala Thr Cys Leu Ser Ile Tyr Tyr Phe Phe Lys Ile Gly
100 105 110

Asn Phe Phe His Pro Leu Phe Leu Trp Met Lys Trp Arg Ile Asp Arg
115 120 125

Val Ile Ser Trp Ile Leu Leu Gly Cys Val Val Leu Ser Val Phe Ile
130 135 140

Ser Leu Pro Ala Thr Glu Asn Leu Asn Ala Asp Phe Arg Phe Cys Val
145 150 155 160

Lys Ala Lys Arg Lys Thr Asn Leu Thr Trp Ser Cys Arg Val Asn Lys
165 170 175

Thr Gln His Ala Ser Thr Lys Leu Phe Leu Asn Leu Ala Thr Leu Leu
180 185 190

Pro Phe Cys Val Cys Leu Met Ser Phe Phe Leu Leu Ile Leu Ser Leu
195 200 205

Arg Arg His Ile Arg Arg Met Gln Leu Ser Ala Thr Gly Cys Arg Asp
210 215 220

Pro Ser Thr Glu Ala His Val Arg Ala Leu Lys Ala Val Ile Ser Phe
225 230 235 240

Leu Leu Leu Phe Ile Ala Tyr Tyr Leu Ser Phe Leu Ile Ala Thr Ser
245 250 255

Ser Tyr Phe Met Pro Glu Xaa Glu Leu Ala Val Ile Phe Gly Glu Ser
260 265 270

Ile Ala Leu Ile Tyr Pro Ser Ser His Ser Phe Ile Leu Ile Leu Gly
275 280 285

Asn Asn Lys Leu Arg His Ala Ser Leu Lys Val Ile Trp Lys Val Met
290 295 300

Ser Ile Leu Lys Gly Arg Lys Phe Gln Gln His Lys Gln Ile
305 310 315

<210> 11
<211> 930
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(930)

<400> 11

```
atg ttc agt cct gca gat aac atc ttt ata atc cta ata act gga gaa       48
Met Phe Ser Pro Ala Asp Asn Ile Phe Ile Ile Leu Ile Thr Gly Glu
1               5                  10                  15

ttc ata cta gga ata ttg ggg aat gga tac att gca cta gtc aac tgg       96
Phe Ile Leu Gly Ile Leu Gly Asn Gly Tyr Ile Ala Leu Val Asn Trp
            20                  25                  30

att gac tgg att aag aag aaa aag att tcc aca gtt gac tac atc ctt      144
Ile Asp Trp Ile Lys Lys Lys Lys Ile Ser Thr Val Asp Tyr Ile Leu
        35                  40                  45

acc aat tta gtt atc gcc aga att tgt ttg atc agt gta atg gtt gta      192
Thr Asn Leu Val Ile Ala Arg Ile Cys Leu Ile Ser Val Met Val Val
    50                  55                  60

aat ggc att gta ata gta ctg aac cca gat gtt tat aca aaa aat aaa      240
Asn Gly Ile Val Ile Val Leu Asn Pro Asp Val Tyr Thr Lys Asn Lys
65                  70                  75                  80

caa cag ata gtc att ttt acc ttc tgg aca ttt gcc aac tac tta aat      288
Gln Gln Ile Val Ile Phe Thr Phe Trp Thr Phe Ala Asn Tyr Leu Asn
                85                  90                  95

atg tgg att acc acc tgc ctt aat gtc ttc tat ttt ctg aag ata gcc      336
Met Trp Ile Thr Thr Cys Leu Asn Val Phe Tyr Phe Leu Lys Ile Ala
            100                 105                 110

agt tcc tct cat cca ctt ttt ctc tgg ctg aag tgg aaa att gat atg      384
Ser Ser Ser His Pro Leu Phe Leu Trp Leu Lys Trp Lys Ile Asp Met
        115                 120                 125

gtg gtg cac tgg atc ctg ctg gga tgc ttt gcc att tcc ttg ttg gtc      432
Val Val His Trp Ile Leu Leu Gly Cys Phe Ala Ile Ser Leu Leu Val
    130                 135                 140

agc ctt ata gca gca ata gta ctg agt tgt gat tat agg ttt cat gca      480
Ser Leu Ile Ala Ala Ile Val Leu Ser Cys Asp Tyr Arg Phe His Ala
145                 150                 155                 160

att gcc aaa cat aaa rga aac att act gaa atg ttc cat gtg agt aaa      528
Ile Ala Lys His Lys Xaa Asn Ile Thr Glu Met Phe His Val Ser Lys
                165                 170                 175

ata cca tac ttt gaa ccc ttr act ctc ttt aac ctg ttt gca att gtc      576
Ile Pro Tyr Phe Glu Pro Xaa Thr Leu Phe Asn Leu Phe Ala Ile Val
            180                 185                 190

cca ttt att gtg tca ctg ata tca ttt ttc ctt tta gta aga tct tta      624
Pro Phe Ile Val Ser Leu Ile Ser Phe Phe Leu Leu Val Arg Ser Leu
```

```
        195                 200                 205

tgg aga cat acc aag caa ata aaa ctc tat gct acc ggc agt aga gac      672
Trp Arg His Thr Lys Gln Ile Lys Leu Tyr Ala Thr Gly Ser Arg Asp
    210                 215                 220

ccc agc aca gaa gtt cat gtg aga gcc att aaa act atg act tca ttt      720
Pro Ser Thr Glu Val His Val Arg Ala Ile Lys Thr Met Thr Ser Phe
225                 230                 235                 240

atc ttc ttt ttt ttc cta tac tat att tct tct att ttg atg acc ttt      768
Ile Phe Phe Phe Phe Leu Tyr Tyr Ile Ser Ser Ile Leu Met Thr Phe
                245                 250                 255

agc tat ctt atg aca aaa tac aag tta gct gtg gag ttt gga gag att      816
Ser Tyr Leu Met Thr Lys Tyr Lys Leu Ala Val Glu Phe Gly Glu Ile
            260                 265                 270

gca gca att ctc yac ccc ttg ggt cac tca ctt att tta att gtt tta      864
Ala Ala Ile Leu Xaa Pro Leu Gly His Ser Leu Ile Leu Ile Val Leu
        275                 280                 285

aat aat aaa ctg agg cag aca ttt gtc aga atg ctg aca tgt aga aaa      912
Asn Asn Lys Leu Arg Gln Thr Phe Val Arg Met Leu Thr Cys Arg Lys
    290                 295                 300

att gcc tgc rtg ata tga                                             930
Ile Ala Cys Xaa Ile
305
```

<210> 12
<211> 309
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (166)..(166)
<223> The 'Xaa' at location 166 stands for Gly, or Arg.

<220>
<221> misc_feature
<222> (183)..(183)
<223> The 'Xaa' at location 183 stands for Leu.

<220>
<221> misc_feature
<222> (277)..(277)
<223> The 'Xaa' at location 277 stands for His, or Tyr.

<220>
<221> misc_feature
<222> (308)..(308)
<223> The 'Xaa' at location 308 stands for Val, or Met.

<400> 12

```
Met Phe Ser Pro Ala Asp Asn Ile Phe Ile Ile Leu Ile Thr Gly Glu
1               5                  10                  15
```

```
Phe Ile Leu Gly Ile Leu Gly Asn Gly Tyr Ile Ala Leu Val Asn Trp
            20                  25                  30

Ile Asp Trp Ile Lys Lys Lys Lys Ile Ser Thr Val Asp Tyr Ile Leu
        35                  40                  45

Thr Asn Leu Val Ile Ala Arg Ile Cys Leu Ile Ser Val Met Val Val
    50                  55                  60

Asn Gly Ile Val Ile Val Leu Asn Pro Asp Val Tyr Thr Lys Asn Lys
65                  70                  75                  80

Gln Gln Ile Val Ile Phe Thr Phe Trp Thr Phe Ala Asn Tyr Leu Asn
                85                  90                  95

Met Trp Ile Thr Thr Cys Leu Asn Val Phe Tyr Phe Leu Lys Ile Ala
            100                 105                 110

Ser Ser Ser His Pro Leu Phe Leu Trp Leu Lys Trp Lys Ile Asp Met
        115                 120                 125

Val Val His Trp Ile Leu Leu Gly Cys Phe Ala Ile Ser Leu Leu Val
    130                 135                 140

Ser Leu Ile Ala Ala Ile Val Leu Ser Cys Asp Tyr Arg Phe His Ala
145                 150                 155                 160

Ile Ala Lys His Lys Xaa Asn Ile Thr Glu Met Phe His Val Ser Lys
                165                 170                 175

Ile Pro Tyr Phe Glu Pro Xaa Thr Leu Phe Asn Leu Phe Ala Ile Val
            180                 185                 190

Pro Phe Ile Val Ser Leu Ile Ser Phe Phe Leu Leu Val Arg Ser Leu
        195                 200                 205

Trp Arg His Thr Lys Gln Ile Lys Leu Tyr Ala Thr Gly Ser Arg Asp
    210                 215                 220

Pro Ser Thr Glu Val His Val Arg Ala Ile Lys Thr Met Thr Ser Phe
225                 230                 235                 240

Ile Phe Phe Phe Phe Leu Tyr Tyr Ile Ser Ser Ile Leu Met Thr Phe
                245                 250                 255
```

```
Ser Tyr Leu Met Thr Lys Tyr Lys Leu Ala Val Glu Phe Gly Glu Ile
            260                 265                 270

Ala Ala Ile Leu Xaa Pro Leu Gly His Ser Leu Ile Leu Ile Val Leu
        275                 280                 285

Asn Asn Lys Leu Arg Gln Thr Phe Val Arg Met Leu Thr Cys Arg Lys
    290                 295                 300

Ile Ala Cys Xaa Ile
305
```

<210> 13
<211> 939
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(939)

<400> 13

```
atg cca agt gca ata gag gca ata tat att att tta att gct ggt gaa        48
Met Pro Ser Ala Ile Glu Ala Ile Tyr Ile Ile Leu Ile Ala Gly Glu
1               5                   10                  15

ttg acc ata ggg att tgg gga aat gga ttc att gta cta gtt aac tgc        96
Leu Thr Ile Gly Ile Trp Gly Asn Gly Phe Ile Val Leu Val Asn Cys
            20                  25                  30

att gac tgg ctc aaa aga aga gat att tcc ttg att gac atc atc ctg       144
Ile Asp Trp Leu Lys Arg Arg Asp Ile Ser Leu Ile Asp Ile Ile Leu
        35                  40                  45

atc agc ttg gcc atc tcc aga atc tgt ctg ctg tgt gta ata tca tta       192
Ile Ser Leu Ala Ile Ser Arg Ile Cys Leu Leu Cys Val Ile Ser Leu
    50                  55                  60

gat ggc ttm ttt atg ctg ctc ttt cca ggt aca tat ggc aat agc gtg       240
Asp Gly Xaa Phe Met Leu Leu Phe Pro Gly Thr Tyr Gly Asn Ser Val
65                  70                  75                  80

cta gta agc att gtg aat gtt gtc tgg aca ttt gcc aat aat tca agt       288
Leu Val Ser Ile Val Asn Val Val Trp Thr Phe Ala Asn Asn Ser Ser
                85                  90                  95

ctc tgg ttt act tct tgc ctc agt atc ttc tat tta ctc aag ata gcc       336
Leu Trp Phe Thr Ser Cys Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala
            100                 105                 110

aat ata tcg cac cca ttt ttc ttc tgg ctg aag cta aag atc aac aag       384
Asn Ile Ser His Pro Phe Phe Phe Trp Leu Lys Leu Lys Ile Asn Lys
        115                 120                 125
```

```
gtc atg ctt gcg att ctt ctg ggg tcc ttt ctt atc tct tta att att      432
Val Met Leu Ala Ile Leu Leu Gly Ser Phe Leu Ile Ser Leu Ile Ile
    130                 135                 140

agt gtt cca aag aat gaw gat atg tgg tat cac ctt ttc aaa gtc agt      480
Ser Val Pro Lys Asn Xaa Asp Met Trp Tyr His Leu Phe Lys Val Ser
145                 150                 155                 160

cat gaa gaa aac att act tgg aaa ttc aaa gtg agt aaa att cca ggt      528
His Glu Glu Asn Ile Thr Trp Lys Phe Lys Val Ser Lys Ile Pro Gly
                165                 170                 175

act ttc aaa cag tta acc ctg aac ctg ggg gyg atg gtt ccc ttt atc      576
Thr Phe Lys Gln Leu Thr Leu Asn Leu Gly Xaa Met Val Pro Phe Ile
            180                 185                 190

ctt tgc ctg atc tca ttt ttc ttg tta ctt ttc tcc cta gtt aga cac      624
Leu Cys Leu Ile Ser Phe Phe Leu Leu Leu Phe Ser Leu Val Arg His
        195                 200                 205

acc aag cag att cga ctg cat gct aca ggg ttc aga gac ccc agt aca      672
Thr Lys Gln Ile Arg Leu His Ala Thr Gly Phe Arg Asp Pro Ser Thr
    210                 215                 220

gag gcc cac atg agg gcc ata aag gca gtg atc atc ttt ctg ctc ctc      720
Glu Ala His Met Arg Ala Ile Lys Ala Val Ile Ile Phe Leu Leu Leu
225                 230                 235                 240

ctc atc gtg tac tac cca gtc ttt ctt gtt atg acc tct agc gct ctg      768
Leu Ile Val Tyr Tyr Pro Val Phe Leu Val Met Thr Ser Ser Ala Leu
                245                 250                 255

att cct cag gga aaa tta gtg ttg atg att ggt gac ata gta act gtc      816
Ile Pro Gln Gly Lys Leu Val Leu Met Ile Gly Asp Ile Val Thr Val
            260                 265                 270

att ttc cca tca agc cat tca ttc att cta att atg gga aat agc aag      864
Ile Phe Pro Ser Ser His Ser Phe Ile Leu Ile Met Gly Asn Ser Lys
        275                 280                 285

ttk agg gaa gct ttt mtg aag atg tta aga ttt gtg aag tgt ttc ctt      912
Xaa Arg Glu Ala Phe Xaa Lys Met Leu Arg Phe Val Lys Cys Phe Leu
    290                 295                 300

aga aga aga aag cct ttt gtt cca tag                                  939
Arg Arg Arg Lys Pro Phe Val Pro
305                 310
```

<210> 14
<211> 312
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (67)..(67)
<223> The 'Xaa' at location 67 stands for Leu, or Phe.

<220>
<221> misc_feature

<222> (150)..(150)
<223> The 'Xaa' at location 150 stands for Glu, or Asp.

<220>
<221> misc_feature
<222> (187)..(187)
<223> The 'Xaa' at location 187 stands for Ala, or Val.

<220>
<221> misc_feature
<222> (289)..(289)
<223> The 'Xaa' at location 289 stands for Leu, or Phe.

<220>
<221> misc_feature
<222> (294)..(294)
<223> The 'Xaa' at location 294 stands for Met, or Leu.

<400> 14

Met Pro Ser Ala Ile Glu Ala Ile Tyr Ile Ile Leu Ile Ala Gly Glu
1 5 10 15

Leu Thr Ile Gly Ile Trp Gly Asn Gly Phe Ile Val Leu Val Asn Cys
20 25 30

Ile Asp Trp Leu Lys Arg Arg Asp Ile Ser Leu Ile Asp Ile Ile Leu
35 40 45

Ile Ser Leu Ala Ile Ser Arg Ile Cys Leu Leu Cys Val Ile Ser Leu
50 55 60

Asp Gly Xaa Phe Met Leu Leu Phe Pro Gly Thr Tyr Gly Asn Ser Val
65 70 75 80

Leu Val Ser Ile Val Asn Val Val Trp Thr Phe Ala Asn Asn Ser Ser
85 90 95

Leu Trp Phe Thr Ser Cys Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala
100 105 110

Asn Ile Ser His Pro Phe Phe Phe Trp Leu Lys Leu Lys Ile Asn Lys
115 120 125

Val Met Leu Ala Ile Leu Leu Gly Ser Phe Leu Ile Ser Leu Ile Ile
130 135 140

Ser Val Pro Lys Asn Xaa Asp Met Trp Tyr His Leu Phe Lys Val Ser
145 150 155 160

His Glu Glu Asn Ile Thr Trp Lys Phe Lys Val Ser Lys Ile Pro Gly

```
                165                 170                 175

Thr Phe Lys Gln Leu Thr Leu Asn Leu Gly Xaa Met Val Pro Phe Ile
            180                 185                 190

Leu Cys Leu Ile Ser Phe Phe Leu Leu Leu Phe Ser Leu Val Arg His
        195                 200                 205

Thr Lys Gln Ile Arg Leu His Ala Thr Gly Phe Arg Asp Pro Ser Thr
    210                 215                 220

Glu Ala His Met Arg Ala Ile Lys Ala Val Ile Ile Phe Leu Leu Leu
225                 230                 235                 240

Leu Ile Val Tyr Tyr Pro Val Phe Leu Val Met Thr Ser Ser Ala Leu
                245                 250                 255

Ile Pro Gln Gly Lys Leu Val Leu Met Ile Gly Asp Ile Val Thr Val
            260                 265                 270

Ile Phe Pro Ser Ser His Ser Phe Ile Leu Ile Met Gly Asn Ser Lys
        275                 280                 285

Xaa Arg Glu Ala Phe Xaa Lys Met Leu Arg Phe Val Lys Cys Phe Leu
    290                 295                 300

Arg Arg Arg Lys Pro Phe Val Pro
305                 310
```

<210> 15
<211> 924
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(924)

<400> 15

```
atg cta cgt gta gtg gaa ggc atc ttc att ttt gtt gta gtt agt gag       48
Met Leu Arg Val Val Glu Gly Ile Phe Ile Phe Val Val Val Ser Glu
1               5                   10                  15

tca gtg ttt ggg gtt ttg ggg aat gga ttt att gga ctt gta aac tgc       96
Ser Val Phe Gly Val Leu Gly Asn Gly Phe Ile Gly Leu Val Asn Cys
            20                  25                  30

att gac tgt gcc aag aat aag ttr tct acg att ggc ttt att ctc acc      144
Ile Asp Cys Ala Lys Asn Lys Xaa Ser Thr Ile Gly Phe Ile Leu Thr
```

```
        35                  40                  45

ggc tta gct att tca aga att ttt ctg ata tgg ata ata att aca gat      192
Gly Leu Ala Ile Ser Arg Ile Phe Leu Ile Trp Ile Ile Ile Thr Asp
    50                  55                  60

gga ttt ata cag ata ttc tct cca aat ata tat gcc tcc ggt aac cta      240
Gly Phe Ile Gln Ile Phe Ser Pro Asn Ile Tyr Ala Ser Gly Asn Leu
65                  70                  75                  80

att gaa tat att agt tac ttt tgg gta att ggt aat caa tca agt atg      288
Ile Glu Tyr Ile Ser Tyr Phe Trp Val Ile Gly Asn Gln Ser Ser Met
                85                  90                  95

tgg ttt gcc acc agc ctc agc atc ttc tat ttc ctg aag ata gca aat      336
Trp Phe Ala Thr Ser Leu Ser Ile Phe Tyr Phe Leu Lys Ile Ala Asn
            100                 105                 110

ttt tcc aac tac ata ttt ctc tgg ttg aag agc aga aca aat atg gtt      384
Phe Ser Asn Tyr Ile Phe Leu Trp Leu Lys Ser Arg Thr Asn Met Val
        115                 120                 125

ctt ccc ttc atg ata gta ttc tta ctt att tca tcg tta ctt aat ttt      432
Leu Pro Phe Met Ile Val Phe Leu Leu Ile Ser Ser Leu Leu Asn Phe
    130                 135                 140

gca tac att gcg aag att ctt aat gat tat aaa ayg aag aat gac aca      480
Ala Tyr Ile Ala Lys Ile Leu Asn Asp Tyr Lys Xaa Lys Asn Asp Thr
145                 150                 155                 160

gtc tgg gat ctc aac atg tat aaa agt gaa tac ttt att ama cag att      528
Val Trp Asp Leu Asn Met Tyr Lys Ser Glu Tyr Phe Ile Xaa Gln Ile
                165                 170                 175

ttg cta aat ctg gga gtc att ttc ttc ttt aca ctr tcc cta att aca      576
Leu Leu Asn Leu Gly Val Ile Phe Phe Phe Thr Xaa Ser Leu Ile Thr
            180                 185                 190

tgt att ttt tta atc att tcc ctt tgg aga cac aac agg cag atg caa      624
Cys Ile Phe Leu Ile Ile Ser Leu Trp Arg His Asn Arg Gln Met Gln
        195                 200                 205

tcr aat gtg aca gga ttg aga gac tcc aac aca gaa gct cat gtg aag      672
Xaa Asn Val Thr Gly Leu Arg Asp Ser Asn Thr Glu Ala His Val Lys
    210                 215                 220

gca atg aaa gtt ttg ata tct ttc atc atc ctc ttt atc ttg tat ttt      720
Ala Met Lys Val Leu Ile Ser Phe Ile Ile Leu Phe Ile Leu Tyr Phe
225                 230                 235                 240

ata ggc atg gcc ata gaa ata tca tgt ttt act gtg cga gaa aac aaa      768
Ile Gly Met Ala Ile Glu Ile Ser Cys Phe Thr Val Arg Glu Asn Lys
                245                 250                 255

ctg ctg ctt atg ttt gga atg aca acc aca gcc atc tat ccc tgg ggt      816
Leu Leu Leu Met Phe Gly Met Thr Thr Thr Ala Ile Tyr Pro Trp Gly
            260                 265                 270

cac tca ttt atc tta att cta gga aac agc aag cta aag caa gcc tct      864
His Ser Phe Ile Leu Ile Leu Gly Asn Ser Lys Leu Lys Gln Ala Ser
        275                 280                 285
```

```
ttg agg gta ctg cag caa ttg aag tgc tgt gag aaa agg aaa aat ctc      912
Leu Arg Val Leu Gln Gln Leu Lys Cys Cys Glu Lys Arg Lys Asn Leu
    290                 295                 300

aga gtc aca tag                                                     924
Arg Val Thr
305
```

<210> 16
<211> 307
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (40)..(40)
<223> The 'Xaa' at location 40 stands for Leu.

<220>
<221> misc_feature
<222> (156)..(156)
<223> The 'Xaa' at location 156 stands for Thr, or Met.

<220>
<221> misc_feature
<222> (174)..(174)
<223> The 'Xaa' at location 174 stands for Lys, or Thr.

<220>
<221> misc_feature
<222> (188)..(188)
<223> The 'Xaa' at location 188 stands for Leu.

<220>
<221> misc_feature
<222> (209)..(209)
<223> The 'Xaa' at location 209 stands for Ser.

<400> 16

```
Met Leu Arg Val Val Glu Gly Ile Phe Ile Phe Val Val Val Ser Glu
1               5                   10                  15

Ser Val Phe Gly Val Leu Gly Asn Gly Phe Ile Gly Leu Val Asn Cys
            20                  25                  30

Ile Asp Cys Ala Lys Asn Lys Xaa Ser Thr Ile Gly Phe Ile Leu Thr
        35                  40                  45

Gly Leu Ala Ile Ser Arg Ile Phe Leu Ile Trp Ile Ile Ile Thr Asp
    50                  55                  60

Gly Phe Ile Gln Ile Phe Ser Pro Asn Ile Tyr Ala Ser Gly Asn Leu
65                  70                  75                  80
```

Ile Glu Tyr Ile Ser Tyr Phe Trp Val Ile Gly Asn Gln Ser Ser Met
85 90 95

Trp Phe Ala Thr Ser Leu Ser Ile Phe Tyr Phe Leu Lys Ile Ala Asn
100 105 110

Phe Ser Asn Tyr Ile Phe Leu Trp Leu Lys Ser Arg Thr Asn Met Val
115 120 125

Leu Pro Phe Met Ile Val Phe Leu Leu Ile Ser Ser Leu Leu Asn Phe
130 135 140

Ala Tyr Ile Ala Lys Ile Leu Asn Asp Tyr Lys Xaa Lys Asn Asp Thr
145 150 155 160

Val Trp Asp Leu Asn Met Tyr Lys Ser Glu Tyr Phe Ile Xaa Gln Ile
165 170 175

Leu Leu Asn Leu Gly Val Ile Phe Phe Phe Thr Xaa Ser Leu Ile Thr
180 185 190

Cys Ile Phe Leu Ile Ile Ser Leu Trp Arg His Asn Arg Gln Met Gln
195 200 205

Xaa Asn Val Thr Gly Leu Arg Asp Ser Asn Thr Glu Ala His Val Lys
210 215 220

Ala Met Lys Val Leu Ile Ser Phe Ile Ile Leu Phe Ile Leu Tyr Phe
225 230 235 240

Ile Gly Met Ala Ile Glu Ile Ser Cys Phe Thr Val Arg Glu Asn Lys
245 250 255

Leu Leu Leu Met Phe Gly Met Thr Thr Thr Ala Ile Tyr Pro Trp Gly
260 265 270

His Ser Phe Ile Leu Ile Leu Gly Asn Ser Lys Leu Lys Gln Ala Ser
275 280 285

Leu Arg Val Leu Gln Gln Leu Lys Cys Cys Glu Lys Arg Lys Asn Leu
290 295 300

Arg Val Thr
305

<210> 17
<211> 912
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(912)

<400> 17

```
atg gaa agt gcc ctg ccg agt atc ttc act ctt gta ata att gca gaa       48
Met Glu Ser Ala Leu Pro Ser Ile Phe Thr Leu Val Ile Ile Ala Glu
1               5                  10                  15

ttc ata att ggg aat ttg agc aat gga ttt ata gta ctg atc aac tgc       96
Phe Ile Ile Gly Asn Leu Ser Asn Gly Phe Ile Val Leu Ile Asn Cys
            20                  25                  30

att gac tgg gtc agt aaa aga gag ctg tcc tca gtc gat aaa ctc ctc      144
Ile Asp Trp Val Ser Lys Arg Glu Leu Ser Ser Val Asp Lys Leu Leu
        35                  40                  45

att atc ttg gca atc tcc aga att ggg ctg atc tgg gaa ata tta gta      192
Ile Ile Leu Ala Ile Ser Arg Ile Gly Leu Ile Trp Glu Ile Leu Val
    50                  55                  60

agt tgg ttt tta gct ctg cat tat cta gcc ata ttt gtg tct gga aca      240
Ser Trp Phe Leu Ala Leu His Tyr Leu Ala Ile Phe Val Ser Gly Thr
65                  70                  75                  80

gga tta aga att atg att ttt agc tgg ata gtt tct aat cac ttc aat      288
Gly Leu Arg Ile Met Ile Phe Ser Trp Ile Val Ser Asn His Phe Asn
                85                  90                  95

ctc tgg ctt gct aca atc ttc agc atc ttt tat ttg ctc aaa ata gcg      336
Leu Trp Leu Ala Thr Ile Phe Ser Ile Phe Tyr Leu Leu Lys Ile Ala
            100                 105                 110

agt ttc tct agc cct gct ttt ctc tat ttg aag tgg aga gta aac aaa      384
Ser Phe Ser Ser Pro Ala Phe Leu Tyr Leu Lys Trp Arg Val Asn Lys
        115                 120                 125

gtg att ctg atg ata ctg cta gga acc ttg gtc ttc tta ttt tta aat      432
Val Ile Leu Met Ile Leu Leu Gly Thr Leu Val Phe Leu Phe Leu Asn
    130                 135                 140

ctg ata caa ata aac atg cat ata aaa gac tgg ctg gac cga tat gaa      480
Leu Ile Gln Ile Asn Met His Ile Lys Asp Trp Leu Asp Arg Tyr Glu
145                 150                 155                 160

aga aac aca act tgg aat ttc agt atg agt gac ttt gaa aca ttt tca      528
Arg Asn Thr Thr Trp Asn Phe Ser Met Ser Asp Phe Glu Thr Phe Ser
                165                 170                 175

gtg tcg gtc aaa ttc act atg act atg ttc agt cta aca cca ttt act      576
Val Ser Val Lys Phe Thr Met Thr Met Phe Ser Leu Thr Pro Phe Thr
            180                 185                 190

gtg gcc ttc atc tct ttt ctc ctg tta att ttc tcc ctg cag aaa cat      624
Val Ala Phe Ile Ser Phe Leu Leu Leu Ile Phe Ser Leu Gln Lys His
```

```
        195                 200                 205

ctc cag aaa atg caa ctc aat tac aaa gga cac aga gac ccc agg acc      672
Leu Gln Lys Met Gln Leu Asn Tyr Lys Gly His Arg Asp Pro Arg Thr
    210                 215                 220

aag gtc cat aca aat gcc ttg aaa att gtg atc tca ttc ctt tta ttc      720
Lys Val His Thr Asn Ala Leu Lys Ile Val Ile Ser Phe Leu Leu Phe
225                 230                 235                 240

tat gct agt ttc ttt cta tgt gtt ctc ata tca tgg att tct gag ctg      768
Tyr Ala Ser Phe Phe Leu Cys Val Leu Ile Ser Trp Ile Ser Glu Leu
                245                 250                 255

tat cag arc aca gtg atc tac atg ctt tgt gag acg att gga gtc ttc      816
Tyr Gln Xaa Thr Val Ile Tyr Met Leu Cys Glu Thr Ile Gly Val Phe
            260                 265                 270

tct cct tca agc cac tcc ttt ctt ctg att cta gga aac gct aag tta      864
Ser Pro Ser Ser His Ser Phe Leu Leu Ile Leu Gly Asn Ala Lys Leu
        275                 280                 285

aga cag gcc ttt ctt ttg gtg gca gct aag gta tgg gct aaa cga tga      912
Arg Gln Ala Phe Leu Leu Val Ala Ala Lys Val Trp Ala Lys Arg
    290                 295                 300
```

<210> 18
<211> 303
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (259)..(259)
<223> The 'Xaa' at location 259 stands for Ser, or Asn.

<400> 18

```
Met Glu Ser Ala Leu Pro Ser Ile Phe Thr Leu Val Ile Ile Ala Glu
1               5                   10                  15

Phe Ile Ile Gly Asn Leu Ser Asn Gly Phe Ile Val Leu Ile Asn Cys
            20                  25                  30

Ile Asp Trp Val Ser Lys Arg Glu Leu Ser Ser Val Asp Lys Leu Leu
        35                  40                  45

Ile Ile Leu Ala Ile Ser Arg Ile Gly Leu Ile Trp Glu Ile Leu Val
    50                  55                  60

Ser Trp Phe Leu Ala Leu His Tyr Leu Ala Ile Phe Val Ser Gly Thr
65                  70                  75                  80

Gly Leu Arg Ile Met Ile Phe Ser Trp Ile Val Ser Asn His Phe Asn
                85                  90                  95
```

```
Leu Trp Leu Ala Thr Ile Phe Ser Ile Phe Tyr Leu Leu Lys Ile Ala
            100                 105                 110

Ser Phe Ser Ser Pro Ala Phe Leu Tyr Leu Lys Trp Arg Val Asn Lys
        115                 120                 125

Val Ile Leu Met Ile Leu Leu Gly Thr Leu Val Phe Leu Phe Leu Asn
    130                 135                 140

Leu Ile Gln Ile Asn Met His Ile Lys Asp Trp Leu Asp Arg Tyr Glu
145                 150                 155                 160

Arg Asn Thr Thr Trp Asn Phe Ser Met Ser Asp Phe Glu Thr Phe Ser
                165                 170                 175

Val Ser Val Lys Phe Thr Met Thr Met Phe Ser Leu Thr Pro Phe Thr
            180                 185                 190

Val Ala Phe Ile Ser Phe Leu Leu Leu Ile Phe Ser Leu Gln Lys His
        195                 200                 205

Leu Gln Lys Met Gln Leu Asn Tyr Lys Gly His Arg Asp Pro Arg Thr
    210                 215                 220

Lys Val His Thr Asn Ala Leu Lys Ile Val Ile Ser Phe Leu Leu Phe
225                 230                 235                 240

Tyr Ala Ser Phe Phe Leu Cys Val Leu Ile Ser Trp Ile Ser Glu Leu
                245                 250                 255

Tyr Gln Xaa Thr Val Ile Tyr Met Leu Cys Glu Thr Ile Gly Val Phe
            260                 265                 270

Ser Pro Ser Ser His Ser Phe Leu Leu Ile Leu Gly Asn Ala Lys Leu
        275                 280                 285

Arg Gln Ala Phe Leu Leu Val Ala Ala Lys Val Trp Ala Lys Arg
    290                 295                 300
```

<210> 19
<211> 954
<212> DNA
<213> Homo sapiens

<220>

<221> CDS
<222> (1)..(954)

<400> 19

```
atg ggt ggt gtc ata aag agc ata ttt aca ttc gtt tta att gtg gaa       48
Met Gly Gly Val Ile Lys Ser Ile Phe Thr Phe Val Leu Ile Val Glu
1               5                   10                  15

ttt ata att gga aat tta gga aat agt ttc ata gca ctg gtg aac tgt       96
Phe Ile Ile Gly Asn Leu Gly Asn Ser Phe Ile Ala Leu Val Asn Cys
            20                  25                  30

att gac tgg gtc aag gga aga aag atc tct tcg gtt gat cgg atc ctc      144
Ile Asp Trp Val Lys Gly Arg Lys Ile Ser Ser Val Asp Arg Ile Leu
        35                  40                  45

act gct ttg gca atc tct cga att agc ctg gtt tgg tta ata ttc gga      192
Thr Ala Leu Ala Ile Ser Arg Ile Ser Leu Val Trp Leu Ile Phe Gly
    50                  55                  60

agc tgg tgt gtg tct gtg ttt ttc cca gct tta ttt gcc act gaa aaa      240
Ser Trp Cys Val Ser Val Phe Phe Pro Ala Leu Phe Ala Thr Glu Lys
65                  70                  75                  80

atg ttc aga atg ctt rct aat atc tgg aca gtg atc aat cat ttt agt      288
Met Phe Arg Met Leu Xaa Asn Ile Trp Thr Val Ile Asn His Phe Ser
                85                  90                  95

gtc tgg tta gct aca ggc ctc ggt act ttt tat ttt ctc aag ata gcc      336
Val Trp Leu Ala Thr Gly Leu Gly Thr Phe Tyr Phe Leu Lys Ile Ala
            100                 105                 110

aat ttt tct aac tct att ttt ctc tac cta aag tgg agr gtt aaa aag      384
Asn Phe Ser Asn Ser Ile Phe Leu Tyr Leu Lys Trp Arg Val Lys Lys
        115                 120                 125

gtg gtt ttg gtg ctg ctt ctt gtg act tcg gtc ttc ttg ttt tta aat      432
Val Val Leu Val Leu Leu Leu Val Thr Ser Val Phe Leu Phe Leu Asn
    130                 135                 140

att gca ctg ata aac atc cat ata aat gcc agt atc aat gga tac aga      480
Ile Ala Leu Ile Asn Ile His Ile Asn Ala Ser Ile Asn Gly Tyr Arg
145                 150                 155                 160

aga aac aag act tgc agt tct gat tca agt aac ttt aca cga ttt tcc      528
Arg Asn Lys Thr Cys Ser Ser Asp Ser Ser Asn Phe Thr Arg Phe Ser
                165                 170                 175

agt ctt att gta tta acc agc act gtg ttc att ttc ata ccc ttt act      576
Ser Leu Ile Val Leu Thr Ser Thr Val Phe Ile Phe Ile Pro Phe Thr
            180                 185                 190

ttg tcc ctg gca atg ttt ctt ctc ctc atc ttc tcc atg tgg aaa cat      624
Leu Ser Leu Ala Met Phe Leu Leu Leu Ile Phe Ser Met Trp Lys His
        195                 200                 205

cgc aag aag atg cag cac act gtc aaa ata tcc gga gac gcc agc acc      672
Arg Lys Lys Met Gln His Thr Val Lys Ile Ser Gly Asp Ala Ser Thr
    210                 215                 220

aaa gcc cac aga gga gtt aaa agt gtg atc act ttc ttc cta ctc tat      720
```

```
Lys Ala His Arg Gly Val Lys Ser Val Ile Thr Phe Phe Leu Leu Tyr
225                 230                 235                 240

gcc att ttc tct ctg tct ttt ttc ata tca gtt tgg acc tct gaa agg      768
Ala Ile Phe Ser Leu Ser Phe Phe Ile Ser Val Trp Thr Ser Glu Arg
                245                 250                 255

ttg gag gaa aat cta att att ctt tcc cag gtg atg gga atg gct tat      816
Leu Glu Glu Asn Leu Ile Ile Leu Ser Gln Val Met Gly Met Ala Tyr
            260                 265                 270

cct tca tgt cac tca tgt gtt ctg att ctt gga aac aag aag ctg aga      864
Pro Ser Cys His Ser Cys Val Leu Ile Leu Gly Asn Lys Lys Leu Arg
        275                 280                 285

cag gcc tct ctg tca gtg cta ctg tgg ctg agg tac atg ttc aaa gat      912
Gln Ala Ser Leu Ser Val Leu Leu Trp Leu Arg Tyr Met Phe Lys Asp
    290                 295                 300

ggg gag ccc tca ggt cac aaa gaa ttt aga gaa tca tct tga              954
Gly Glu Pro Ser Gly His Lys Glu Phe Arg Glu Ser Ser
305                 310                 315
```

<210> 20
<211> 317
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (86)..(86)
<223> The 'Xaa' at location 86 stands for Ala, or Thr.

<400> 20

```
Met Gly Gly Val Ile Lys Ser Ile Phe Thr Phe Val Leu Ile Val Glu
1               5                   10                  15

Phe Ile Ile Gly Asn Leu Gly Asn Ser Phe Ile Ala Leu Val Asn Cys
            20                  25                  30

Ile Asp Trp Val Lys Gly Arg Lys Ile Ser Ser Val Asp Arg Ile Leu
        35                  40                  45

Thr Ala Leu Ala Ile Ser Arg Ile Ser Leu Val Trp Leu Ile Phe Gly
    50                  55                  60

Ser Trp Cys Val Ser Val Phe Phe Pro Ala Leu Phe Ala Thr Glu Lys
65                  70                  75                  80

Met Phe Arg Met Leu Xaa Asn Ile Trp Thr Val Ile Asn His Phe Ser
                85                  90                  95

Val Trp Leu Ala Thr Gly Leu Gly Thr Phe Tyr Phe Leu Lys Ile Ala
```

100 105 110

Asn Phe Ser Asn Ser Ile Phe Leu Tyr Leu Lys Trp Arg Val Lys Lys
115 120 125

Val Val Leu Val Leu Leu Leu Val Thr Ser Val Phe Leu Phe Leu Asn
130 135 140

Ile Ala Leu Ile Asn Ile His Ile Asn Ala Ser Ile Asn Gly Tyr Arg
145 150 155 160

Arg Asn Lys Thr Cys Ser Ser Asp Ser Ser Asn Phe Thr Arg Phe Ser
165 170 175

Ser Leu Ile Val Leu Thr Ser Thr Val Phe Ile Phe Ile Pro Phe Thr
180 185 190

Leu Ser Leu Ala Met Phe Leu Leu Leu Ile Phe Ser Met Trp Lys His
195 200 205

Arg Lys Lys Met Gln His Thr Val Lys Ile Ser Gly Asp Ala Ser Thr
210 215 220

Lys Ala His Arg Gly Val Lys Ser Val Ile Thr Phe Phe Leu Leu Tyr
225 230 235 240

Ala Ile Phe Ser Leu Ser Phe Phe Ile Ser Val Trp Thr Ser Glu Arg
245 250 255

Leu Glu Glu Asn Leu Ile Ile Leu Ser Gln Val Met Gly Met Ala Tyr
260 265 270

Pro Ser Cys His Ser Cys Val Leu Ile Leu Gly Asn Lys Lys Leu Arg
275 280 285

Gln Ala Ser Leu Ser Val Leu Leu Trp Leu Arg Tyr Met Phe Lys Asp
290 295 300

Gly Glu Pro Ser Gly His Lys Glu Phe Arg Glu Ser Ser
305 310 315

<210> 21
<211> 876
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(876)

<400> 21
atg ata ccc atc caa ctc act gtc ttc ttc atg atc atc tat gtg ctt 48
Met Ile Pro Ile Gln Leu Thr Val Phe Phe Met Ile Ile Tyr Val Leu
1 5 10 15

gag tcc ttg aca att att gtg cag agc agc cta att gtt gca gtg ctg 96
Glu Ser Leu Thr Ile Ile Val Gln Ser Ser Leu Ile Val Ala Val Leu
20 25 30

ggc aga gaa tgg ctg caa gtc aga agg ctg atg cct gtg gac atg att 144
Gly Arg Glu Trp Leu Gln Val Arg Arg Leu Met Pro Val Asp Met Ile
35 40 45

ctc atc agc ctg ggc atc tct cgc ttc tgt cta cag tgg gca tca atg 192
Leu Ile Ser Leu Gly Ile Ser Arg Phe Cys Leu Gln Trp Ala Ser Met
50 55 60

ctg aac aat ttt tgc tcc tat ttt aat ttg aat tat gta ctt tgc aac 240
Leu Asn Asn Phe Cys Ser Tyr Phe Asn Leu Asn Tyr Val Leu Cys Asn
65 70 75 80

tta aca atc acc tgg gaa ttt ttt aat atc ctt aca ttc tgg tta aac 288
Leu Thr Ile Thr Trp Glu Phe Phe Asn Ile Leu Thr Phe Trp Leu Asn
85 90 95

agc ttg ctt acy rts ttc tac tgc atc aag gtc tct tct ttc acc cat 336
Ser Leu Leu Xaa Xaa Phe Tyr Cys Ile Lys Val Ser Ser Phe Thr His
100 105 110

cac atc ttt ctc tgg ctg agg tgg aga att ttg agg ttg ttt ccc tgg 384
His Ile Phe Leu Trp Leu Arg Trp Arg Ile Leu Arg Leu Phe Pro Trp
115 120 125

ata tta ctg ggt tct ctg atg att act tgt gta aca atc atc cct tca 432
Ile Leu Leu Gly Ser Leu Met Ile Thr Cys Val Thr Ile Ile Pro Ser
130 135 140

gct att ggg aat tac att caa att cag yta ctc acc atg gag cat cta 480
Ala Ile Gly Asn Tyr Ile Gln Ile Gln Xaa Leu Thr Met Glu His Leu
145 150 155 160

cca aga aac agc act gta act gac aaa ctt gaa aak ttt cat cag tat 528
Pro Arg Asn Ser Thr Val Thr Asp Lys Leu Glu Xaa Phe His Gln Tyr
165 170 175

cag ttc cag gct cat aca gtt gca ttg gtt att cct ttc atc ctg ttc 576
Gln Phe Gln Ala His Thr Val Ala Leu Val Ile Pro Phe Ile Leu Phe
180 185 190

ctg gcc tcc acc atc ttt ctc atg gca tca ctg acc aag cag ata caa 624
Leu Ala Ser Thr Ile Phe Leu Met Ala Ser Leu Thr Lys Gln Ile Gln
195 200 205

cat cat agc act ggt cac tgc aat cca agc atg aaa gcg crc ttc act 672
His His Ser Thr Gly His Cys Asn Pro Ser Met Lys Ala Xaa Phe Thr
210 215 220

```
gcc ctg agg tcc ctt gcc gtc tta ttt att gtg ttt acc tct tac ttt      720
Ala Leu Arg Ser Leu Ala Val Leu Phe Ile Val Phe Thr Ser Tyr Phe
225                 230                 235                 240

cta acc ata ctc atc acc att ata ggt act cta ttt gat aag aga tgt      768
Leu Thr Ile Leu Ile Thr Ile Ile Gly Thr Leu Phe Asp Lys Arg Cys
                245                 250                 255

tgg tta tgg gtc tgg gaa gct ttt gtc tat gct ttc atc tta atg cat      816
Trp Leu Trp Val Trp Glu Ala Phe Val Tyr Ala Phe Ile Leu Met His
            260                 265                 270

tcc act tca ctg atg ctg agc agc cct acr ttg aaa agg att cta aag      864
Ser Thr Ser Leu Met Leu Ser Ser Pro Xaa Leu Lys Arg Ile Leu Lys
        275                 280                 285

gga aag tgc tag                                                      876
Gly Lys Cys
    290
```

<210> 22
<211> 291
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (100)..(100)
<223> The 'Xaa' at location 100 stands for Thr.

<220>
<221> misc_feature
<222> (101)..(101)
<223> The 'Xaa' at location 101 stands for Val, Met, or Ile.

<220>
<221> misc_feature
<222> (154)..(154)
<223> The 'Xaa' at location 154 stands for Leu.

<220>
<221> misc_feature
<222> (172)..(172)
<223> The 'Xaa' at location 172 stands for Lys, or Asn.

<220>
<221> misc_feature
<222> (222)..(222)
<223> The 'Xaa' at location 222 stands for Arg, or His.

<220>
<221> misc_feature
<222> (282)..(282)
<223> The 'Xaa' at location 282 stands for Thr.

<400> 22

```
Met Ile Pro Ile Gln Leu Thr Val Phe Phe Met Ile Ile Tyr Val Leu
1               5                   10                  15
```

Glu Ser Leu Thr Ile Ile Val Gln Ser Ser Leu Ile Val Ala Val Leu
20 25 30

Gly Arg Glu Trp Leu Gln Val Arg Arg Leu Met Pro Val Asp Met Ile
35 40 45

Leu Ile Ser Leu Gly Ile Ser Arg Phe Cys Leu Gln Trp Ala Ser Met
50 55 60

Leu Asn Asn Phe Cys Ser Tyr Phe Asn Leu Asn Tyr Val Leu Cys Asn
65 70 75 80

Leu Thr Ile Thr Trp Glu Phe Phe Asn Ile Leu Thr Phe Trp Leu Asn
85 90 95

Ser Leu Leu Xaa Xaa Phe Tyr Cys Ile Lys Val Ser Ser Phe Thr His
100 105 110

His Ile Phe Leu Trp Leu Arg Trp Arg Ile Leu Arg Leu Phe Pro Trp
115 120 125

Ile Leu Leu Gly Ser Leu Met Ile Thr Cys Val Thr Ile Ile Pro Ser
130 135 140

Ala Ile Gly Asn Tyr Ile Gln Ile Gln Xaa Leu Thr Met Glu His Leu
145 150 155 160

Pro Arg Asn Ser Thr Val Thr Asp Lys Leu Glu Xaa Phe His Gln Tyr
165 170 175

Gln Phe Gln Ala His Thr Val Ala Leu Val Ile Pro Phe Ile Leu Phe
180 185 190

Leu Ala Ser Thr Ile Phe Leu Met Ala Ser Leu Thr Lys Gln Ile Gln
195 200 205

His His Ser Thr Gly His Cys Asn Pro Ser Met Lys Ala Xaa Phe Thr
210 215 220

Ala Leu Arg Ser Leu Ala Val Leu Phe Ile Val Phe Thr Ser Tyr Phe
225 230 235 240

Leu Thr Ile Leu Ile Thr Ile Ile Gly Thr Leu Phe Asp Lys Arg Cys
245 250 255

```
Trp Leu Trp Val Trp Glu Ala Phe Val Tyr Ala Phe Ile Leu Met His
            260                 265                 270

Ser Thr Ser Leu Met Leu Ser Ser Pro Xaa Leu Lys Arg Ile Leu Lys
        275                 280                 285

Gly Lys Cys
    290
```

<210> 23
<211> 1002
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1002)

<400> 23

```
atg ttg act cta act cgc atc cgc act gtg tcc tat gaa gtc agg agt       48
Met Leu Thr Leu Thr Arg Ile Arg Thr Val Ser Tyr Glu Val Arg Ser
1               5                   10                  15

aca ttt ctg ttc att tca gtc ctg gag ttt gca gtg ggg ttt ctg acc       96
Thr Phe Leu Phe Ile Ser Val Leu Glu Phe Ala Val Gly Phe Leu Thr
            20                  25                  30

aat gcc ttc gtt ttc ttg gtg aat ttt tgg gat gta gtg aag agg cag      144
Asn Ala Phe Val Phe Leu Val Asn Phe Trp Asp Val Val Lys Arg Gln
        35                  40                  45

sca ctg agc aac agt gat tgt gtg ctg ctg tgt ctc agc atc agc cgg      192
Xaa Leu Ser Asn Ser Asp Cys Val Leu Leu Cys Leu Ser Ile Ser Arg
    50                  55                  60

ctt ttc ctg cat gga ctg ctg ttc ctg agt gct atc cag ctt acc crc      240
Leu Phe Leu His Gly Leu Leu Phe Leu Ser Ala Ile Gln Leu Thr Xaa
65                  70                  75                  80

ttc cag aag ttg agt gaa cca ctg aac cac agc tac caa gcc atc atc      288
Phe Gln Lys Leu Ser Glu Pro Leu Asn His Ser Tyr Gln Ala Ile Ile
                85                  90                  95

atg cta tgg atg att gca aac caa gcc aac ctc tgg ctt gct gcc tgc      336
Met Leu Trp Met Ile Ala Asn Gln Ala Asn Leu Trp Leu Ala Ala Cys
            100                 105                 110

ctc agc ctg ctt tac tgc tcc aag ctc atc cgt ttc tct cac acc ttc      384
Leu Ser Leu Leu Tyr Cys Ser Lys Leu Ile Arg Phe Ser His Thr Phe
        115                 120                 125

ctg atc tgc ttg gca agc tgg gtc tcc agg aag atc tcc cag atg ctc      432
Leu Ile Cys Leu Ala Ser Trp Val Ser Arg Lys Ile Ser Gln Met Leu
    130                 135                 140

ctg ggt att att ctt tgc tcc tgc atc tgc act gtc ctc tgt gtt tgg      480
```

```
Leu Gly Ile Ile Leu Cys Ser Cys Ile Cys Thr Val Leu Cys Val Trp
145                 150                 155                 160

tgc ttt ttt agc aga cct cac ttc aca gtc aca act gtg cta ttc atg      528
Cys Phe Phe Ser Arg Pro His Phe Thr Val Thr Thr Val Leu Phe Met
                165                 170                 175

aat aac aat aca agg ctc aac tgg cag aat aaa gat ctc aat tta ttt      576
Asn Asn Asn Thr Arg Leu Asn Trp Gln Asn Lys Asp Leu Asn Leu Phe
            180                 185                 190

tat tcc ttt ctc ttc tgc tat ctg tgg tct gtg cct cct ttc cta ttg      624
Tyr Ser Phe Leu Phe Cys Tyr Leu Trp Ser Val Pro Pro Phe Leu Leu
        195                 200                 205

ttt ctg gtt tct tct ggg atg ctg act gtc tcc ctg gga agg cac atg      672
Phe Leu Val Ser Ser Gly Met Leu Thr Val Ser Leu Gly Arg His Met
    210                 215                 220

agg aca atg aag gtc tat acc aga aac tct cgt gac ccc agc ctg gag      720
Arg Thr Met Lys Val Tyr Thr Arg Asn Ser Arg Asp Pro Ser Leu Glu
225                 230                 235                 240

gcc cac att aaa gcc ctc aag tct ctt gtc tcc ttt ttc tgc ttc ttt      768
Ala His Ile Lys Ala Leu Lys Ser Leu Val Ser Phe Phe Cys Phe Phe
                245                 250                 255

gtg ata tca tcc tgt gyt gcc ttc atc tct gtg ccc cta ctg att ctg      816
Val Ile Ser Ser Cys Xaa Ala Phe Ile Ser Val Pro Leu Leu Ile Leu
            260                 265                 270

tgg ygc gac aaa ata ggg gtg atg gtt tgt gtt ggg ata atg gca gct      864
Trp Xaa Asp Lys Ile Gly Val Met Val Cys Val Gly Ile Met Ala Ala
        275                 280                 285

tgt ccc tct ggg cat gca gcc rtc ctg atc tca ggc aat gcc aag ttg      912
Cys Pro Ser Gly His Ala Ala Xaa Leu Ile Ser Gly Asn Ala Lys Leu
    290                 295                 300

agg aga gct gtg atg acc att ctg ctc tgg gct cag agc agc ctg aag      960
Arg Arg Ala Val Met Thr Ile Leu Leu Trp Ala Gln Ser Ser Leu Lys
305                 310                 315                 320

gta aga gcc gac cac aag gca gat tcc cgg aca ctg tgc tga              1002
Val Arg Ala Asp His Lys Ala Asp Ser Arg Thr Leu Cys
                325                 330
```

<210> 24
<211> 333
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (49)..(49)
<223> The 'Xaa' at location 49 stands for Ala, or Pro.

<220>
<221> misc_feature
<222> (80)..(80)

<223> The 'Xaa' at location 80 stands for Arg, or His.

<220>
<221> misc_feature
<222> (262)..(262)
<223> The 'Xaa' at location 262 stands for Ala, or Val.

<220>
<221> misc_feature
<222> (274)..(274)
<223> The 'Xaa' at location 274 stands for Arg, or Cys.

<220>
<221> misc_feature
<222> (296)..(296)
<223> The 'Xaa' at location 296 stands for Val, or Ile.

<400> 24

Met Leu Thr Leu Thr Arg Ile Arg Thr Val Ser Tyr Glu Val Arg Ser
1 5 10 15

Thr Phe Leu Phe Ile Ser Val Leu Glu Phe Ala Val Gly Phe Leu Thr
20 25 30

Asn Ala Phe Val Phe Leu Val Asn Phe Trp Asp Val Val Lys Arg Gln
35 40 45

Xaa Leu Ser Asn Ser Asp Cys Val Leu Leu Cys Leu Ser Ile Ser Arg
50 55 60

Leu Phe Leu His Gly Leu Leu Phe Leu Ser Ala Ile Gln Leu Thr Xaa
65 70 75 80

Phe Gln Lys Leu Ser Glu Pro Leu Asn His Ser Tyr Gln Ala Ile Ile
85 90 95

Met Leu Trp Met Ile Ala Asn Gln Ala Asn Leu Trp Leu Ala Ala Cys
100 105 110

Leu Ser Leu Leu Tyr Cys Ser Lys Leu Ile Arg Phe Ser His Thr Phe
115 120 125

Leu Ile Cys Leu Ala Ser Trp Val Ser Arg Lys Ile Ser Gln Met Leu
130 135 140

Leu Gly Ile Ile Leu Cys Ser Cys Ile Cys Thr Val Leu Cys Val Trp
145 150 155 160

Cys Phe Phe Ser Arg Pro His Phe Thr Val Thr Thr Val Leu Phe Met
165 170 175

```
Asn Asn Asn Thr Arg Leu Asn Trp Gln Asn Lys Asp Leu Asn Leu Phe
            180                 185                 190

Tyr Ser Phe Leu Phe Cys Tyr Leu Trp Ser Val Pro Pro Phe Leu Leu
        195                 200                 205

Phe Leu Val Ser Ser Gly Met Leu Thr Val Ser Leu Gly Arg His Met
    210                 215                 220

Arg Thr Met Lys Val Tyr Thr Arg Asn Ser Arg Asp Pro Ser Leu Glu
225                 230                 235                 240

Ala His Ile Lys Ala Leu Lys Ser Leu Val Ser Phe Phe Cys Phe Phe
                245                 250                 255

Val Ile Ser Ser Cys Xaa Ala Phe Ile Ser Val Pro Leu Leu Ile Leu
            260                 265                 270

Trp Xaa Asp Lys Ile Gly Val Met Val Cys Val Gly Ile Met Ala Ala
        275                 280                 285

Cys Pro Ser Gly His Ala Ala Xaa Leu Ile Ser Gly Asn Ala Lys Leu
    290                 295                 300

Arg Arg Ala Val Met Thr Ile Leu Leu Trp Ala Gln Ser Ser Leu Lys
305                 310                 315                 320

Val Arg Ala Asp His Lys Ala Asp Ser Arg Thr Leu Cys
                325                 330
```

<210> 25
<211> 1014
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1014)

<400> 25

```
atg cta ggg aga tgt ttt cct cca gac acc aaa gag aag caa cag ctc       48
Met Leu Gly Arg Cys Phe Pro Pro Asp Thr Lys Glu Lys Gln Gln Leu
1               5                   10                  15

aga atg act aaa ctc tgc gat cct gca gaa agt gaa ttg tcg cca ttt       96
Arg Met Thr Lys Leu Cys Asp Pro Ala Glu Ser Glu Leu Ser Pro Phe
            20                  25                  30
```

```
ctc atc acc tta att tta gca gtt tta ctt gct gaa tac ctc att ggt      144
Leu Ile Thr Leu Ile Leu Ala Val Leu Leu Ala Glu Tyr Leu Ile Gly
        35                  40                  45

atc att gca aat ggt ttc atc atg gct ata cat gca gct gaa tgg gtt      192
Ile Ile Ala Asn Gly Phe Ile Met Ala Ile His Ala Ala Glu Trp Val
    50                  55                  60

caa aat aag gca gtt tcc aca agt ggc agg atc ctg gtt ttc ctg agt      240
Gln Asn Lys Ala Val Ser Thr Ser Gly Arg Ile Leu Val Phe Leu Ser
65                  70                  75                  80

gta tcc aga ata gct ctc caa agc ctc atg atg tta gaa att acc atc      288
Val Ser Arg Ile Ala Leu Gln Ser Leu Met Met Leu Glu Ile Thr Ile
                85                  90                  95

agc tca acc tcc cta agt ttt tat tct gaa gac gct gta tat tat gca      336
Ser Ser Thr Ser Leu Ser Phe Tyr Ser Glu Asp Ala Val Tyr Tyr Ala
            100                 105                 110

ttc aaa ata agt ttt ata ttc tta aat ttt tgt agc ctg tgg ttt gct      384
Phe Lys Ile Ser Phe Ile Phe Leu Asn Phe Cys Ser Leu Trp Phe Ala
        115                 120                 125

gcc tgg ctc agt ttc ttc tac ttt gtg aag att gcc aat ttc tcc tac      432
Ala Trp Leu Ser Phe Phe Tyr Phe Val Lys Ile Ala Asn Phe Ser Tyr
    130                 135                 140

ccc ctt ttc ctc aaa ctg agg tgg aga att act gga ttg ata ccc tgg      480
Pro Leu Phe Leu Lys Leu Arg Trp Arg Ile Thr Gly Leu Ile Pro Trp
145                 150                 155                 160

ctt ctg tgg ctg tcc gtg ttt att tcc ttc agt cac agc atg ttc tgc      528
Leu Leu Trp Leu Ser Val Phe Ile Ser Phe Ser His Ser Met Phe Cys
                165                 170                 175

atc aac atc tgc act gtg tat tgt aac aat tct ttc cct atc cac tcc      576
Ile Asn Ile Cys Thr Val Tyr Cys Asn Asn Ser Phe Pro Ile His Ser
            180                 185                 190

tyc aac tcc act rag aaa aca tac ttg tct gag atc aat gtg gtc ggt      624
Xaa Asn Ser Thr Xaa Lys Thr Tyr Leu Ser Glu Ile Asn Val Val Gly
        195                 200                 205

ctg gct ttt ttc ttt aac ctg ggg att gtg act cct ctg atc atg ttc      672
Leu Ala Phe Phe Phe Asn Leu Gly Ile Val Thr Pro Leu Ile Met Phe
    210                 215                 220

atc ctg aca gcc acc ctg ctg atc ctc tct ctc aag aga cac acc cta      720
Ile Leu Thr Ala Thr Leu Leu Ile Leu Ser Leu Lys Arg His Thr Leu
225                 230                 235                 240

cac atg gga agc aat gcc aca ggg tcc aac gac ccc agc atg gag gct      768
His Met Gly Ser Asn Ala Thr Gly Ser Asn Asp Pro Ser Met Glu Ala
                245                 250                 255

cac atg ggg gcc atc aaa gct atc agc tac ttt ctc att ctc tac att      816
His Met Gly Ala Ile Lys Ala Ile Ser Tyr Phe Leu Ile Leu Tyr Ile
            260                 265                 270

ttc aat gca gtt gct ctg ttt atc tac ctg tcc aac atg ttt gac atc      864
```

```
Phe Asn Ala Val Ala Leu Phe Ile Tyr Leu Ser Asn Met Phe Asp Ile
        275                 280                 285

aac agt ctg tgg aat aat ttg tgc cag atc atc atg gct gcc tac cct      912
Asn Ser Leu Trp Asn Asn Leu Cys Gln Ile Ile Met Ala Ala Tyr Pro
    290                 295                 300

gcc agc cac tca att cta ctg att caa gat aac cct ggg ctg aga aga      960
Ala Ser His Ser Ile Leu Leu Ile Gln Asp Asn Pro Gly Leu Arg Arg
305                 310                 315                 320

gcc tgg agc ggc ttc agc ttc gac ttc atc ttt acc caa aag agt gga     1008
Ala Trp Ser Gly Phe Ser Phe Asp Phe Ile Phe Thr Gln Lys Ser Gly
                325                 330                 335

ctc tga                                                            1014
Leu
```

<210> 26
<211> 337
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (193)..(193)
<223> The 'Xaa' at location 193 stands for Ser, or Phe.

<220>
<221> misc_feature
<222> (197)..(197)
<223> The 'Xaa' at location 197 stands for Glu, or Lys.

<400> 26

```
Met Leu Gly Arg Cys Phe Pro Pro Asp Thr Lys Glu Lys Gln Gln Leu
1               5                   10                  15

Arg Met Thr Lys Leu Cys Asp Pro Ala Glu Ser Glu Leu Ser Pro Phe
            20                  25                  30

Leu Ile Thr Leu Ile Leu Ala Val Leu Leu Ala Glu Tyr Leu Ile Gly
        35                  40                  45

Ile Ile Ala Asn Gly Phe Ile Met Ala Ile His Ala Ala Glu Trp Val
    50                  55                  60

Gln Asn Lys Ala Val Ser Thr Ser Gly Arg Ile Leu Val Phe Leu Ser
65                  70                  75                  80

Val Ser Arg Ile Ala Leu Gln Ser Leu Met Met Leu Glu Ile Thr Ile
                85                  90                  95
```

Ser Ser Thr Ser Leu Ser Phe Tyr Ser Glu Asp Ala Val Tyr Tyr Ala
100 105 110

Phe Lys Ile Ser Phe Ile Phe Leu Asn Phe Cys Ser Leu Trp Phe Ala
115 120 125

Ala Trp Leu Ser Phe Phe Tyr Phe Val Lys Ile Ala Asn Phe Ser Tyr
130 135 140

Pro Leu Phe Leu Lys Leu Arg Trp Arg Ile Thr Gly Leu Ile Pro Trp
145 150 155 160

Leu Leu Trp Leu Ser Val Phe Ile Ser Phe Ser His Ser Met Phe Cys
165 170 175

Ile Asn Ile Cys Thr Val Tyr Cys Asn Asn Ser Phe Pro Ile His Ser
180 185 190

Xaa Asn Ser Thr Xaa Lys Thr Tyr Leu Ser Glu Ile Asn Val Val Gly
195 200 205

Leu Ala Phe Phe Phe Asn Leu Gly Ile Val Thr Pro Leu Ile Met Phe
210 215 220

Ile Leu Thr Ala Thr Leu Leu Ile Leu Ser Leu Lys Arg His Thr Leu
225 230 235 240

His Met Gly Ser Asn Ala Thr Gly Ser Asn Asp Pro Ser Met Glu Ala
245 250 255

His Met Gly Ala Ile Lys Ala Ile Ser Tyr Phe Leu Ile Leu Tyr Ile
260 265 270

Phe Asn Ala Val Ala Leu Phe Ile Tyr Leu Ser Asn Met Phe Asp Ile
275 280 285

Asn Ser Leu Trp Asn Asn Leu Cys Gln Ile Ile Met Ala Ala Tyr Pro
290 295 300

Ala Ser His Ser Ile Leu Leu Ile Gln Asp Asn Pro Gly Leu Arg Arg
305 310 315 320

Ala Trp Ser Gly Phe Ser Phe Asp Phe Ile Phe Thr Gln Lys Ser Gly
325 330 335

Leu

<210> 27
<211> 972
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(972)

<400> 27

```
atg gca acg gtg aac aca gat gcc aca gat aaa gac ata tcc aag ttc        48
Met Ala Thr Val Asn Thr Asp Ala Thr Asp Lys Asp Ile Ser Lys Phe
1               5                  10                  15

aag gtc acc ttc act ttg gtg gtc tcc gga ata gag tgc atc act ggc        96
Lys Val Thr Phe Thr Leu Val Val Ser Gly Ile Glu Cys Ile Thr Gly
            20                  25                  30

atc ctt ggg agt ggc ttc atc acg gcc atc tat ggg gct gag tgg gcc       144
Ile Leu Gly Ser Gly Phe Ile Thr Ala Ile Tyr Gly Ala Glu Trp Ala
        35                  40                  45

agg ggc aaa aca ctc ccc act ggt gac cgc att atg ttg atg ctg agc       192
Arg Gly Lys Thr Leu Pro Thr Gly Asp Arg Ile Met Leu Met Leu Ser
    50                  55                  60

ttt tcc agg ctc ttg cta cag att tgg atg atg ctg gag aac att ttc       240
Phe Ser Arg Leu Leu Leu Gln Ile Trp Met Met Leu Glu Asn Ile Phe
65                  70                  75                  80

agt ctg cta ttc cga att gtt tat aac caa aac tca gtg tat atc ctc       288
Ser Leu Leu Phe Arg Ile Val Tyr Asn Gln Asn Ser Val Tyr Ile Leu
                85                  90                  95

ttc aaa gtc atc act gtc ttt ctg aac cat tcc aat ctc tgg ttt gct       336
Phe Lys Val Ile Thr Val Phe Leu Asn His Ser Asn Leu Trp Phe Ala
            100                 105                 110

gcc tgg ctc aaa gtc ttc tat tgt ctt aga att gca aac ttc aat cat       384
Ala Trp Leu Lys Val Phe Tyr Cys Leu Arg Ile Ala Asn Phe Asn His
        115                 120                 125


cct ttg ttc ttc ctg atg aag agg aaa atc ata gtg ctg atg cct tgg       432
Pro Leu Phe Phe Leu Met Lys Arg Lys Ile Ile Val Leu Met Pro Trp
    130                 135                 140

ctt ctc agg ctg tca gtg ttg gtt tcc tta agc ttc agc ttt cct ctc       480
Leu Leu Arg Leu Ser Val Leu Val Ser Leu Ser Phe Ser Phe Pro Leu
145                 150                 155                 160

tcg aga gat gtc ttc aat gtg tat gtg aat agc tcc att cct atc ccc       528
Ser Arg Asp Val Phe Asn Val Tyr Val Asn Ser Ser Ile Pro Ile Pro
                165                 170                 175

tcc tcc aac tcc acg gag aag aag tac ttc tmt gag acc aat atg gtc       576
Ser Ser Asn Ser Thr Glu Lys Lys Tyr Phe Xaa Glu Thr Asn Met Val
```

```
            180                 185                 190

aac ctg gta ttt ttc tat aac atg ggg atc ttc gtt cct ctg atc atg      624
Asn Leu Val Phe Phe Tyr Asn Met Gly Ile Phe Val Pro Leu Ile Met
        195                 200                 205

ttc atc ctg gca gcc acc ctg ctg atc ctc tct ctc aag aga cac acc      672
Phe Ile Leu Ala Ala Thr Leu Leu Ile Leu Ser Leu Lys Arg His Thr
    210                 215                 220

cta cac atg gga agc aat gcc aca ggg tcc agg gac ccc agc atg aag      720
Leu His Met Gly Ser Asn Ala Thr Gly Ser Arg Asp Pro Ser Met Lys
225                 230                 235                 240

gct cac ata ggg gcc atc aaa gcc acc agc tac ttt ctc atc ctc tac      768
Ala His Ile Gly Ala Ile Lys Ala Thr Ser Tyr Phe Leu Ile Leu Tyr
                245                 250                 255


att ttc aat gca att gct cta ttt ctt tcc acg tcc aac atc ttt gac      816
Ile Phe Asn Ala Ile Ala Leu Phe Leu Ser Thr Ser Asn Ile Phe Asp
            260                 265                 270

rct tac agt tcc tgg aat att ttg tgc aag atc atc atg gct gcc tac      864
Xaa Tyr Ser Ser Trp Asn Ile Leu Cys Lys Ile Ile Met Ala Ala Tyr
        275                 280                 285

cct gcc ggc cac tca gta caa ctg atc ttg ggc aac cct ggg ctg aga      912
Pro Ala Gly His Ser Val Gln Leu Ile Leu Gly Asn Pro Gly Leu Arg
    290                 295                 300

aga gcc tgg aag cgg ttt cag cac caa gtt cct ctt tac cta aaa ggg      960
Arg Ala Trp Lys Arg Phe Gln His Gln Val Pro Leu Tyr Leu Lys Gly
305                 310                 315                 320

cag act ctg tga                                                     972
Gln Thr Leu
```

<210> 28
<211> 323
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (187)..(187)
<223> The 'Xaa' at location 187 stands for Tyr, or Ser.

<220>
<221> misc_feature
<222> (273)..(273)
<223> The 'Xaa' at location 273 stands for Ala, or Thr.

<400> 28

```
Met Ala Thr Val Asn Thr Asp Ala Thr Asp Lys Asp Ile Ser Lys Phe
1               5                   10                  15
```

Lys Val Thr Phe Thr Leu Val Val Ser Gly Ile Glu Cys Ile Thr Gly
20 25 30

Ile Leu Gly Ser Gly Phe Ile Thr Ala Ile Tyr Gly Ala Glu Trp Ala
35 40 45

Arg Gly Lys Thr Leu Pro Thr Gly Asp Arg Ile Met Leu Met Leu Ser
50 55 60

Phe Ser Arg Leu Leu Leu Gln Ile Trp Met Met Leu Glu Asn Ile Phe
65 70 75 80

Ser Leu Leu Phe Arg Ile Val Tyr Asn Gln Asn Ser Val Tyr Ile Leu
85 90 95

Phe Lys Val Ile Thr Val Phe Leu Asn His Ser Asn Leu Trp Phe Ala
100 105 110

Ala Trp Leu Lys Val Phe Tyr Cys Leu Arg Ile Ala Asn Phe Asn His
115 120 125

Pro Leu Phe Phe Leu Met Lys Arg Lys Ile Ile Val Leu Met Pro Trp
130 135 140

Leu Leu Arg Leu Ser Val Leu Val Ser Leu Ser Phe Ser Phe Pro Leu
145 150 155 160

Ser Arg Asp Val Phe Asn Val Tyr Val Asn Ser Ser Ile Pro Ile Pro
165 170 175

Ser Ser Asn Ser Thr Glu Lys Lys Tyr Phe Xaa Glu Thr Asn Met Val
180 185 190

Asn Leu Val Phe Phe Tyr Asn Met Gly Ile Phe Val Pro Leu Ile Met
195 200 205

Phe Ile Leu Ala Ala Thr Leu Leu Ile Leu Ser Leu Lys Arg His Thr
210 215 220

Leu His Met Gly Ser Asn Ala Thr Gly Ser Arg Asp Pro Ser Met Lys
225 230 235 240

Ala His Ile Gly Ala Ile Lys Ala Thr Ser Tyr Phe Leu Ile Leu Tyr
245 250 255

Ile Phe Asn Ala Ile Ala Leu Phe Leu Ser Thr Ser Asn Ile Phe Asp

```
            260                 265                 270

Xaa Tyr Ser Ser Trp Asn Ile Leu Cys Lys Ile Ile Met Ala Ala Tyr
        275                 280                 285

Pro Ala Gly His Ser Val Gln Leu Ile Leu Gly Asn Pro Gly Leu Arg
    290                 295                 300

Arg Ala Trp Lys Arg Phe Gln His Gln Val Pro Leu Tyr Leu Lys Gly
305                 310                 315                 320

Gln Thr Leu
```

<210> 29
<211> 924
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(924)

<400> 29

```
atg caa gca gca ctg acg gcc ttc ttc gtg ttg ctc ttt agc ctg ctg      48
Met Gln Ala Ala Leu Thr Ala Phe Phe Val Leu Leu Phe Ser Leu Leu
1               5                   10                  15

agt ctt ctg ggg att gca gcg aat ggc ttc att gtg ctg gtg ctg ggc      96
Ser Leu Leu Gly Ile Ala Ala Asn Gly Phe Ile Val Leu Val Leu Gly
            20                  25                  30

agg gag tgg ctg cga tat ggc agg ttg ctg ccc ttg gat atg atc ctc     144
Arg Glu Trp Leu Arg Tyr Gly Arg Leu Leu Pro Leu Asp Met Ile Leu
        35                  40                  45

att agc ttg ggt gcc tcc cgc ttc tgc ctg cag ttg gtt ggg acr gtg     192
Ile Ser Leu Gly Ala Ser Arg Phe Cys Leu Gln Leu Val Gly Xaa Val
    50                  55                  60

cac aac ttc tac tac tct gcc cag aag gtc gag tac tct ggg ggt ctc     240
His Asn Phe Tyr Tyr Ser Ala Gln Lys Val Glu Tyr Ser Gly Gly Leu
65                  70                  75                  80

ggc cga cag ttc ttc cat cta cac tgg cac ttc ctg aac tca gcc acc     288
Gly Arg Gln Phe Phe His Leu His Trp His Phe Leu Asn Ser Ala Thr
                85                  90                  95

ttc tgg ttt tgc agc tgg ctc agt gtc ctg ttc tgt gtg aag att gct     336
Phe Trp Phe Cys Ser Trp Leu Ser Val Leu Phe Cys Val Lys Ile Ala
            100                 105                 110

aac atc aca cac tcc acc ttc ctg tgg ctg aag tgg agg ttc cya ggg     384
Asn Ile Thr His Ser Thr Phe Leu Trp Leu Lys Trp Arg Phe Xaa Gly
        115                 120                 125
```

```
tgg gtg ccc tgg ctc ctg ttg ggc tct gtc ctg atc tcc ttc atc ata      432
Trp Val Pro Trp Leu Leu Leu Gly Ser Val Leu Ile Ser Phe Ile Ile
    130                 135                 140

acc ctg ctg ttt ttt tgg gtg aac tac cct gta tat caa gaa ttt tta      480
Thr Leu Leu Phe Phe Trp Val Asn Tyr Pro Val Tyr Gln Glu Phe Leu
145                 150                 155                 160

att aga aaa ttt tct ggg aac atg acc tac aag tgg aat aca agg ata      528
Ile Arg Lys Phe Ser Gly Asn Met Thr Tyr Lys Trp Asn Thr Arg Ile
                165                 170                 175

gaa aca tac tat ttc cca tcc ctg aaa ctg gtc atc tgg tca att cct      576
Glu Thr Tyr Tyr Phe Pro Ser Leu Lys Leu Val Ile Trp Ser Ile Pro
            180                 185                 190

ttt tct gwt ttt ctg gtc tca att atg ctg tta att aat tct ctg agg      624
Phe Ser Xaa Phe Leu Val Ser Ile Met Leu Leu Ile Asn Ser Leu Arg
        195                 200                 205

agg cat act cag aga atg cag cac aac ggg cac agc ctg cag gac ccc      672
Arg His Thr Gln Arg Met Gln His Asn Gly His Ser Leu Gln Asp Pro
    210                 215                 220

agc acc cag gct cac acc aga gct ctg aag tcc ctc atc tcc ttc ctc      720
Ser Thr Gln Ala His Thr Arg Ala Leu Lys Ser Leu Ile Ser Phe Leu
225                 230                 235                 240

att ctt tat gct ctg tcc ttt ctg tcc ctg atc att gat gcc gca aaa      768
Ile Leu Tyr Ala Leu Ser Phe Leu Ser Leu Ile Ile Asp Ala Ala Lys
                245                 250                 255

ttt atc tcc atg cag aac gac ttt tac tgg cca tgg caa att gca gtc      816
Phe Ile Ser Met Gln Asn Asp Phe Tyr Trp Pro Trp Gln Ile Ala Val
            260                 265                 270

tac ctg tgc ata tct gtc cat ccc ttc atc ctc atc ttc agc aac ctc      864
Tyr Leu Cys Ile Ser Val His Pro Phe Ile Leu Ile Phe Ser Asn Leu
        275                 280                 285

aag ctt cga agc gtg ttc tca cag ctc ctg ttg ttg gca agg ggc ttc      912
Lys Leu Arg Ser Val Phe Ser Gln Leu Leu Leu Leu Ala Arg Gly Phe
    290                 295                 300

tgg gtg gcc tga                                                      924
Trp Val Ala
305
```

<210> 30
<211> 307
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (63)..(63)
<223> The 'Xaa' at location 63 stands for Thr.

<220>

<221> misc_feature
<222> (127)..(127)
<223> The 'Xaa' at location 127 stands for Pro, or Leu.

<220>
<221> misc_feature
<222> (195)..(195)
<223> The 'Xaa' at location 195 stands for Asp, or Val.

<400> 30

Met Gln Ala Ala Leu Thr Ala Phe Phe Val Leu Leu Phe Ser Leu Leu
1 5 10 15

Ser Leu Leu Gly Ile Ala Ala Asn Gly Phe Ile Val Leu Val Leu Gly
20 25 30

Arg Glu Trp Leu Arg Tyr Gly Arg Leu Leu Pro Leu Asp Met Ile Leu
35 40 45

Ile Ser Leu Gly Ala Ser Arg Phe Cys Leu Gln Leu Val Gly Xaa Val
50 55 60

His Asn Phe Tyr Tyr Ser Ala Gln Lys Val Glu Tyr Ser Gly Gly Leu
65 70 75 80

Gly Arg Gln Phe Phe His Leu His Trp His Phe Leu Asn Ser Ala Thr
85 90 95

Phe Trp Phe Cys Ser Trp Leu Ser Val Leu Phe Cys Val Lys Ile Ala
100 105 110

Asn Ile Thr His Ser Thr Phe Leu Trp Leu Lys Trp Arg Phe Xaa Gly
115 120 125

Trp Val Pro Trp Leu Leu Leu Gly Ser Val Leu Ile Ser Phe Ile Ile
130 135 140

Thr Leu Leu Phe Phe Trp Val Asn Tyr Pro Val Tyr Gln Glu Phe Leu
145 150 155 160

Ile Arg Lys Phe Ser Gly Asn Met Thr Tyr Lys Trp Asn Thr Arg Ile
165 170 175

Glu Thr Tyr Tyr Phe Pro Ser Leu Lys Leu Val Ile Trp Ser Ile Pro
180 185 190

Phe Ser Xaa Phe Leu Val Ser Ile Met Leu Leu Ile Asn Ser Leu Arg
195 200 205

```
Arg His Thr Gln Arg Met Gln His Asn Gly His Ser Leu Gln Asp Pro
    210                 215                 220

Ser Thr Gln Ala His Thr Arg Ala Leu Lys Ser Leu Ile Ser Phe Leu
225                 230                 235                 240

Ile Leu Tyr Ala Leu Ser Phe Leu Ser Leu Ile Ile Asp Ala Ala Lys
                245                 250                 255

Phe Ile Ser Met Gln Asn Asp Phe Tyr Trp Pro Trp Gln Ile Ala Val
            260                 265                 270

Tyr Leu Cys Ile Ser Val His Pro Phe Ile Leu Ile Phe Ser Asn Leu
        275                 280                 285

Lys Leu Arg Ser Val Phe Ser Gln Leu Leu Leu Leu Ala Arg Gly Phe
    290                 295                 300

Trp Val Ala
305
```

<210> 31
<211> 930
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(930)

<400> 31

```
atg ata act ttt cta ccc atc att ttt tcc agt ctg gta gtg gtt aca      48
Met Ile Thr Phe Leu Pro Ile Ile Phe Ser Ser Leu Val Val Val Thr
1               5                   10                  15

ttt gtt att gga aat ttt gct aat ggc ttc ata gca ctg gta aat tcc      96
Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30

att gag tsg ttc aag aga caa aag atc tcc ttt gct gac caa att ctc     144
Ile Glu Xaa Phe Lys Arg Gln Lys Ile Ser Phe Ala Asp Gln Ile Leu
        35                  40                  45

act gct ctg gcg gtc tcc aga gtt ggt ttg ctc tgg gta tta tta tta     192
Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Leu Leu
    50                  55                  60

aac tgg tat tca act gtg ttg aat cca gct ttt aat agt gta gaa gta     240
Asn Trp Tyr Ser Thr Val Leu Asn Pro Ala Phe Asn Ser Val Glu Val
65                  70                  75                  80
```

```
aga act act gct tat aat atc tgg gca gtr atc aac cat ttc agc aac      288
Arg Thr Thr Ala Tyr Asn Ile Trp Ala Xaa Ile Asn His Phe Ser Asn
                85                  90                  95

tgg ctt gct act acc ctc agc ata ttt tat ttg ctc aag att gcc aat      336
Trp Leu Ala Thr Thr Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
            100                 105                 110

ttc tcc aac ttt att ttt ctt cac tta aag agg aga gtt aag agt gtc      384
Phe Ser Asn Phe Ile Phe Leu His Leu Lys Arg Arg Val Lys Ser Val
        115                 120                 125

att ctg gtg atg ttg ttg ggg cct ttg cta ttt ttg gct tgt cat ctt      432
Ile Leu Val Met Leu Leu Gly Pro Leu Leu Phe Leu Ala Cys His Leu
    130                 135                 140

ttt gtg ata aac atg aat gag att gtg sgg aca aaa gaa ttt gaa gga      480
Phe Val Ile Asn Met Asn Glu Ile Val Xaa Thr Lys Glu Phe Glu Gly
145                 150                 155                 160

aac atg act tgg aag atc aaa ttg aag agk gca atg tac ttt tca aat      528
Asn Met Thr Trp Lys Ile Lys Leu Lys Xaa Ala Met Tyr Phe Ser Asn
                165                 170                 175


atg act gta acc atg gta gca aac tta gta ccc ttc act ctg acc cta      576
Met Thr Val Thr Met Val Ala Asn Leu Val Pro Phe Thr Leu Thr Leu
            180                 185                 190

cta tct ttt atg ctg tta atc tkt tct ttg tgt aaa cat ctc aag aag      624
Leu Ser Phe Met Leu Leu Ile Xaa Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

atg cag ctc crt ggt aaa gga tct caa gat ccc agc acs aag gtc cac      672
Met Gln Leu Xaa Gly Lys Gly Ser Gln Asp Pro Ser Xaa Lys Val His
    210                 215                 220

ata aaa gct ttg caa act gtg atc tcc ttc ctc ttg tta tgt gcc att      720
Ile Lys Ala Leu Gln Thr Val Ile Ser Phe Leu Leu Leu Cys Ala Ile
225                 230                 235                 240

tac ttt ctg tcc ata atg ata tca gtt tgg agt ttt gga agt ctg gaa      768
Tyr Phe Leu Ser Ile Met Ile Ser Val Trp Ser Phe Gly Ser Leu Glu
                245                 250                 255

aac aaa cct gtc ttc atg ttc tgc aaa gct att aga ttc agc tat cct      816
Asn Lys Pro Val Phe Met Phe Cys Lys Ala Ile Arg Phe Ser Tyr Pro
            260                 265                 270

tca atc cac cca ttc atc ctg att tgg gga aac aag aag cta aag cag      864
Ser Ile His Pro Phe Ile Leu Ile Trp Gly Asn Lys Lys Leu Lys Gln
        275                 280                 285

act ttt ctt tca gtt ttk ygg caa rtg agg tac tgg gtg aaa gga gag      912
Thr Phe Leu Ser Val Xaa Xaa Gln Xaa Arg Tyr Trp Val Lys Gly Glu
    290                 295                 300


aag act tca tct cca tga                                             930
Lys Thr Ser Ser Pro
```

305

<210> 32
<211> 309
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (35)..(35)
<223> The 'Xaa' at location 35 stands for Trp, or Ser.

<220>
<221> misc_feature
<222> (90)..(90)
<223> The 'Xaa' at location 90 stands for Val.

<220>
<221> misc_feature
<222> (154)..(154)
<223> The 'Xaa' at location 154 stands for Gly, or Arg.

<220>
<221> misc_feature
<222> (170)..(170)
<223> The 'Xaa' at location 170 stands for Arg, or Ser.

<220>
<221> misc_feature
<222> (200)..(200)
<223> The 'Xaa' at location 200 stands for Cys, or Phe.

<220>
<221> misc_feature
<222> (212)..(212)
<223> The 'Xaa' at location 212 stands for Arg, or His.

<220>
<221> misc_feature
<222> (221)..(221)
<223> The 'Xaa' at location 221 stands for Thr.

<220>
<221> misc_feature
<222> (294)..(294)
<223> The 'Xaa' at location 294 stands for Leu, or Phe.

<220>
<221> misc_feature
<222> (295)..(295)
<223> The 'Xaa' at location 295 stands for Arg, or Trp.

<220>
<221> misc_feature
<222> (297)..(297)
<223> The 'Xaa' at location 297 stands for Val, or Met.

<400> 32

Met Ile Thr Phe Leu Pro Ile Ile Phe Ser Ser Leu Val Val Val Thr

1 5 10 15

Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
20 25 30

Ile Glu Xaa Phe Lys Arg Gln Lys Ile Ser Phe Ala Asp Gln Ile Leu
35 40 45

Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Leu Leu
50 55 60

Asn Trp Tyr Ser Thr Val Leu Asn Pro Ala Phe Asn Ser Val Glu Val
65 70 75 80

Arg Thr Thr Ala Tyr Asn Ile Trp Ala Xaa Ile Asn His Phe Ser Asn
85 90 95

Trp Leu Ala Thr Thr Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
100 105 110

Phe Ser Asn Phe Ile Phe Leu His Leu Lys Arg Arg Val Lys Ser Val
115 120 125

Ile Leu Val Met Leu Leu Gly Pro Leu Leu Phe Leu Ala Cys His Leu
130 135 140

Phe Val Ile Asn Met Asn Glu Ile Val Xaa Thr Lys Glu Phe Glu Gly
145 150 155 160

Asn Met Thr Trp Lys Ile Lys Leu Lys Xaa Ala Met Tyr Phe Ser Asn
165 170 175

Met Thr Val Thr Met Val Ala Asn Leu Val Pro Phe Thr Leu Thr Leu
180 185 190

Leu Ser Phe Met Leu Leu Ile Xaa Ser Leu Cys Lys His Leu Lys Lys
195 200 205

Met Gln Leu Xaa Gly Lys Gly Ser Gln Asp Pro Ser Xaa Lys Val His
210 215 220

Ile Lys Ala Leu Gln Thr Val Ile Ser Phe Leu Leu Leu Cys Ala Ile
225 230 235 240

Tyr Phe Leu Ser Ile Met Ile Ser Val Trp Ser Phe Gly Ser Leu Glu
245 250 255

```
Asn Lys Pro Val Phe Met Phe Cys Lys Ala Ile Arg Phe Ser Tyr Pro
            260                 265                 270

Ser Ile His Pro Phe Ile Leu Ile Trp Gly Asn Lys Lys Leu Lys Gln
        275                 280                 285

Thr Phe Leu Ser Val Xaa Xaa Gln Xaa Arg Tyr Trp Val Lys Gly Glu
    290                 295                 300

Lys Thr Ser Ser Pro
305
```

<210> 33
<211> 930
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(930)

<400> 33

```
atg aca act ttt ata ccc atc att ttt tcc agt gtg gta gtg gtt cta       48
Met Thr Thr Phe Ile Pro Ile Ile Phe Ser Ser Val Val Val Val Leu
1               5                   10                  15

ttt gtt att gga aat ttt gct aat ggc ttc ata gca ttg gta aat tcc       96
Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30

att gag ygg gtc aag aga caa aag atc tct ttt gct gac cag att ctc      144
Ile Glu Xaa Val Lys Arg Gln Lys Ile Ser Phe Ala Asp Gln Ile Leu
        35                  40                  45

act gct ctg gcg gtc tcc aga gtt ggt ttg ctc tgg gta tta tta tta      192
Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Leu Leu
    50                  55                  60

aat tgg tat tca act gtg ttt aat cca gct ttt tat agt gta gaa gta      240
Asn Trp Tyr Ser Thr Val Phe Asn Pro Ala Phe Tyr Ser Val Glu Val
65                  70                  75                  80

aga act act gct tat aat gtc tgg gca gta acc ggc cat ttc agc aac      288
Arg Thr Thr Ala Tyr Asn Val Trp Ala Val Thr Gly His Phe Ser Asn

                85                  90                  95

tgg ctt gct act agc ctc agc ata ttt tat ttg ctc aag att gcc aat      336
Trp Leu Ala Thr Ser Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
            100                 105                 110

ttc tcc aac ctt att ttt ctt cac tta aag agg aga gtt aag agt gtc      384
Phe Ser Asn Leu Ile Phe Leu His Leu Lys Arg Arg Val Lys Ser Val
        115                 120                 125
```

```
att ctg gtg atg ctg ttg ggg cct tta cta ttt ttg gcy tgt caa ctt      432
Ile Leu Val Met Leu Leu Gly Pro Leu Leu Phe Leu Ala Cys Gln Leu
    130                 135                 140

ttt gtg ata aac atg aaa gag att gta cgg aca aaa gaa tat gaa gga      480
Phe Val Ile Asn Met Lys Glu Ile Val Arg Thr Lys Glu Tyr Glu Gly
145                 150                 155                 160

aac wtg act tgg aag atc aaa ttg agg agt gca gtg tac ctt tca gat      528
Asn Xaa Thr Trp Lys Ile Lys Leu Arg Ser Ala Val Tyr Leu Ser Asp
                165                 170                 175

gcg act gta acc acg cta gga aac tta gtg ccc ttc act ctg acc ctg      576
Ala Thr Val Thr Thr Leu Gly Asn Leu Val Pro Phe Thr Leu Thr Leu
            180                 185                 190

cta tgt ttt ttg ctg tta atc trt tct ctg tgt aaa cat ctc aag aag      624
Leu Cys Phe Leu Leu Leu Ile Xaa Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

atg cag ctc cat ggt aaa gga tct saa gat ccc agc acc aag gtc cac      672
Met Gln Leu His Gly Lys Gly Ser Xaa Asp Pro Ser Thr Lys Val His

    210                 215                 220

ata aaa gyt ttg caa act gtg atc ttt ttc ctc ttg tta tgt gcc rtt      720
Ile Lys Xaa Leu Gln Thr Val Ile Phe Phe Leu Leu Leu Cys Ala Xaa
225                 230                 235                 240

tac ttt ctg tcc ata atg ata tcr gtt tgg agt ttt ggg agt ctg gaa      768
Tyr Phe Leu Ser Ile Met Ile Xaa Val Trp Ser Phe Gly Ser Leu Glu
                245                 250                 255

aac aaa cct gtc ttc atg ttc tgc aaa gct att aga ttc agc tat cct      816
Asn Lys Pro Val Phe Met Phe Cys Lys Ala Ile Arg Phe Ser Tyr Pro
            260                 265                 270

tca atc cac csa ttc atc ctg att tgr gga aac aag aag cta aag cag      864
Ser Ile His Xaa Phe Ile Leu Ile Xaa Gly Asn Lys Lys Leu Lys Gln
        275                 280                 285

act ttt ctt tca gtt ttg cgg caa gtg agg tac tgg gtg aaa gga gag      912
Thr Phe Leu Ser Val Leu Arg Gln Val Arg Tyr Trp Val Lys Gly Glu
    290                 295                 300

aag cct tca tct cca tga                                             930
Lys Pro Ser Ser Pro
305
```

<210> 34
<211> 309
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (35)..(35)
<223> The 'Xaa' at location 35 stands for Arg, or Trp.

<220>
<221> misc_feature
<222> (162)..(162)
<223> The 'Xaa' at location 162 stands for Met, or Leu.

<220>
<221> misc_feature
<222> (200)..(200)
<223> The 'Xaa' at location 200 stands for Cys, or Tyr.

<220>
<221> misc_feature
<222> (217)..(217)
<223> The 'Xaa' at location 217 stands for Glu, or Gln.

<220>
<221> misc_feature
<222> (227)..(227)
<223> The 'Xaa' at location 227 stands for Ala, or Val.

<220>
<221> misc_feature
<222> (240)..(240)
<223> The 'Xaa' at location 240 stands for Val, or Ile.

<220>
<221> misc_feature
<222> (248)..(248)
<223> The 'Xaa' at location 248 stands for Ser.

<220>
<221> misc_feature
<222> (276)..(276)
<223> The 'Xaa' at location 276 stands for Arg, or Pro.

<220>
<221> misc_feature
<222> (281)..(281)
<223> The 'Xaa' at location 281 stands for Trp or STOP.

<400> 34

```
Met Thr Thr Phe Ile Pro Ile Ile Phe Ser Ser Val Val Val Val Leu
1               5                  10                  15

Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30

Ile Glu Xaa Val Lys Arg Gln Lys Ile Ser Phe Ala Asp Gln Ile Leu
        35                  40                  45

Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Leu Leu
    50                  55                  60

Asn Trp Tyr Ser Thr Val Phe Asn Pro Ala Phe Tyr Ser Val Glu Val
65                  70                  75                  80
```

Arg Thr Thr Ala Tyr Asn Val Trp Ala Val Thr Gly His Phe Ser Asn
85 90 95

Trp Leu Ala Thr Ser Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
100 105 110

Phe Ser Asn Leu Ile Phe Leu His Leu Lys Arg Arg Val Lys Ser Val
115 120 125

Ile Leu Val Met Leu Leu Gly Pro Leu Leu Phe Leu Ala Cys Gln Leu
130 135 140

Phe Val Ile Asn Met Lys Glu Ile Val Arg Thr Lys Glu Tyr Glu Gly
145 150 155 160

Asn Xaa Thr Trp Lys Ile Lys Leu Arg Ser Ala Val Tyr Leu Ser Asp
165 170 175

Ala Thr Val Thr Thr Leu Gly Asn Leu Val Pro Phe Thr Leu Thr Leu
180 185 190

Leu Cys Phe Leu Leu Leu Ile Xaa Ser Leu Cys Lys His Leu Lys Lys
195 200 205

Met Gln Leu His Gly Lys Gly Ser Xaa Asp Pro Ser Thr Lys Val His
210 215 220

Ile Lys Xaa Leu Gln Thr Val Ile Phe Phe Leu Leu Leu Cys Ala Xaa
225 230 235 240

Tyr Phe Leu Ser Ile Met Ile Xaa Val Trp Ser Phe Gly Ser Leu Glu
245 250 255

Asn Lys Pro Val Phe Met Phe Cys Lys Ala Ile Arg Phe Ser Tyr Pro
260 265 270

Ser Ile His Xaa Phe Ile Leu Ile Xaa Gly Asn Lys Lys Leu Lys Gln
275 280 285

Thr Phe Leu Ser Val Leu Arg Gln Val Arg Tyr Trp Val Lys Gly Glu
290 295 300

Lys Pro Ser Ser Pro
305

<210> 35
<211> 900
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(900)

<400> 35

```
atg ata act ttt ctg ccc atc att ttt tcc att cta ata gtg gtt aca       48
Met Ile Thr Phe Leu Pro Ile Ile Phe Ser Ile Leu Ile Val Val Thr
1               5                  10                  15

ttt gtg att gga aat ttt gct aat ggc ttc ata gca ttg gta aat tcc       96
Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30

att gag tgg ktt aag aga caa aag atc tct ttt gct gac caa att ctc      144
Ile Glu Trp Xaa Lys Arg Gln Lys Ile Ser Phe Ala Asp Gln Ile Leu
        35                  40                  45

act gct ctg gca gtc tcc aga gtt ggt tta ctc tgg gta tta gta tta      192
Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Val Leu
    50                  55                  60

aat tgg tat gca act gag ttg aat cca gct ttt aac agt ata gaa gta      240
Asn Trp Tyr Ala Thr Glu Leu Asn Pro Ala Phe Asn Ser Ile Glu Val
65                  70                  75                  80

aga att act gct tac aat gtc tgg gca gta atc aac cat ttc agc aac      288
Arg Ile Thr Ala Tyr Asn Val Trp Ala Val Ile Asn His Phe Ser Asn
                85                  90                  95

tgg ctt gct act agc ctc agc ata ttt tat ttg ctc aag att gcc aat      336
Trp Leu Ala Thr Ser Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
            100                 105                 110

ttc tcc aac ctt att ttt ctt cac tta aag agg aga gtt aag agt gtt      384
Phe Ser Asn Leu Ile Phe Leu His Leu Lys Arg Arg Val Lys Ser Val
        115                 120                 125

gtt ctg gtg ata cta ttg ggg cct ttg cta ttt ttg gtt tgt cat ctt      432
Val Leu Val Ile Leu Leu Gly Pro Leu Leu Phe Leu Val Cys His Leu
    130                 135                 140

ttt gtg ata aac atg aat cag att ata tgg aca aaa gaa tat gaa gga      480
Phe Val Ile Asn Met Asn Gln Ile Ile Trp Thr Lys Glu Tyr Glu Gly
145                 150                 155                 160

aac atg act tgg aag atc aaa ctg agg agt gca atg tac ctt tca aat      528
Asn Met Thr Trp Lys Ile Lys Leu Arg Ser Ala Met Tyr Leu Ser Asn
                165                 170                 175

aca acg gta acc atc cta gca aac tta gtt ccc ttc act ctg acc ctg      576
Thr Thr Val Thr Ile Leu Ala Asn Leu Val Pro Phe Thr Leu Thr Leu
            180                 185                 190

ata tct ttt ctg ctg tta atc tgt tct ctg tgt aaa cat ctc aaa aag      624
Ile Ser Phe Leu Leu Leu Ile Cys Ser Leu Cys Lys His Leu Lys Lys
```

```
        195                 200                 205

atg cag ctc cat ggc aaa gga tct caa gat ccc agc atg aag gtc cac      672
Met Gln Leu His Gly Lys Gly Ser Gln Asp Pro Ser Met Lys Val His
    210                 215                 220

ata aaa gct wtg caa act gtg acc tcc ttc ctc ttg tta tgt gcc att      720
Ile Lys Ala Xaa Gln Thr Val Thr Ser Phe Leu Leu Leu Cys Ala Ile
225                 230                 235                 240

tac ttt ctg tcc ata atc atg tca gtt trg agt ttt gag agt ctg gaa      768
Tyr Phe Leu Ser Ile Ile Met Ser Val Xaa Ser Phe Glu Ser Leu Glu
                245                 250                 255

aac aaa cct gtc ttc atg ttc tgc gaa gct att gca ttc agc tat cct      816
Asn Lys Pro Val Phe Met Phe Cys Glu Ala Ile Ala Phe Ser Tyr Pro
            260                 265                 270

tca acc cac cca ttc atc ctg att tgg gga aac aag aag cta aag yag      864
Ser Thr His Pro Phe Ile Leu Ile Trp Gly Asn Lys Lys Leu Lys Xaa
        275                 280                 285

act ttt ctt tca gtt ttg tgg caa atg agg tac tga                      900
Thr Phe Leu Ser Val Leu Trp Gln Met Arg Tyr
    290                 295
```

<210> 36
<211> 299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (36)..(36)
<223> The 'Xaa' at location 36 stands for Val, or Phe.

<220>
<221> misc_feature
<222> (228)..(228)
<223> The 'Xaa' at location 228 stands for Met, or Leu.

<220>
<221> misc_feature
<222> (250)..(250)
<223> The 'Xaa' at location 250 stands for Trp, or STOP

<220>
<221> misc_feature
<222> (288)..(288)
<223> The 'Xaa' at location 288 stands for Gln, or STOP

<400> 36

```
Met Ile Thr Phe Leu Pro Ile Ile Phe Ser Ile Leu Ile Val Val Thr
1               5                   10                  15


Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30
```

Ile Glu Trp Xaa Lys Arg Gln Lys Ile Ser Phe Ala Asp Gln Ile Leu
35 40 45

Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Val Leu
50 55 60

Asn Trp Tyr Ala Thr Glu Leu Asn Pro Ala Phe Asn Ser Ile Glu Val
65 70 75 80

Arg Ile Thr Ala Tyr Asn Val Trp Ala Val Ile Asn His Phe Ser Asn
85 90 95

Trp Leu Ala Thr Ser Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
100 105 110

Phe Ser Asn Leu Ile Phe Leu His Leu Lys Arg Arg Val Lys Ser Val
115 120 125

Val Leu Val Ile Leu Leu Gly Pro Leu Leu Phe Leu Val Cys His Leu
130 135 140

Phe Val Ile Asn Met Asn Gln Ile Ile Trp Thr Lys Glu Tyr Glu Gly
145 150 155 160

Asn Met Thr Trp Lys Ile Lys Leu Arg Ser Ala Met Tyr Leu Ser Asn
165 170 175

Thr Thr Val Thr Ile Leu Ala Asn Leu Val Pro Phe Thr Leu Thr Leu
180 185 190

Ile Ser Phe Leu Leu Leu Ile Cys Ser Leu Cys Lys His Leu Lys Lys
195 200 205

Met Gln Leu His Gly Lys Gly Ser Gln Asp Pro Ser Met Lys Val His
210 215 220

Ile Lys Ala Xaa Gln Thr Val Thr Ser Phe Leu Leu Leu Cys Ala Ile
225 230 235 240

Tyr Phe Leu Ser Ile Ile Met Ser Val Xaa Ser Phe Glu Ser Leu Glu
245 250 255

Asn Lys Pro Val Phe Met Phe Cys Glu Ala Ile Ala Phe Ser Tyr Pro
260 265 270

```
Ser Thr His Pro Phe Ile Leu Ile Trp Gly Asn Lys Lys Leu Lys Xaa
        275                 280                 285

Thr Phe Leu Ser Val Leu Trp Gln Met Arg Tyr
    290                 295
```

<210> 37
<211> 929
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(927)

<400> 37

```
atg ata act ttt ctg ccc atc att ttt tcc att cta ata gtg gtt ata       48
Met Ile Thr Phe Leu Pro Ile Ile Phe Ser Ile Leu Ile Val Val Ile
1               5                   10                  15

ttt gtt att gga aat ttt gct aat ggc ttc ata gca ttg gta aat tcc       96
Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30

att gag tgg gtc aag aga caa aag atc tcc ttt gtt gac caa att ctc      144
Ile Glu Trp Val Lys Arg Gln Lys Ile Ser Phe Val Asp Gln Ile Leu
        35                  40                  45

act gct ctg gcg gtc tcc aga gtt ggt ttg ctc tgg gtg tta tta cta      192
Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Leu Leu
    50                  55                  60

cat tgg tat gca act cag ttg aat cca gct ttt tat agt gta gaa gta      240
His Trp Tyr Ala Thr Gln Leu Asn Pro Ala Phe Tyr Ser Val Glu Val
65                  70                  75                  80

aga att act gct tat aat gtc tgg gca gta acc aac cat ttc agc agc      288
Arg Ile Thr Ala Tyr Asn Val Trp Ala Val Thr Asn His Phe Ser Ser
                85                  90                  95

tgg ctt gct act agc ctc agc atg ttt tat ttg ctc agg att gcc aat      336
Trp Leu Ala Thr Ser Leu Ser Met Phe Tyr Leu Leu Arg Ile Ala Asn
            100                 105                 110

ttc tcc aac ctt att ttt ctt cgc ata aag agg aga gtt aag agt gtt      384
Phe Ser Asn Leu Ile Phe Leu Arg Ile Lys Arg Arg Val Lys Ser Val
        115                 120                 125

gtt ctg gtg ata ctg ttg ggg cct ttg cta ttt ttg gtt tgt cat ctt      432
Val Leu Val Ile Leu Leu Gly Pro Leu Leu Phe Leu Val Cys His Leu
    130                 135                 140

ttt gtg ata aac atg gat gag act gta tgg aca aaa gaa tat gaa gga      480
Phe Val Ile Asn Met Asp Glu Thr Val Trp Thr Lys Glu Tyr Glu Gly
145                 150                 155                 160

aac gtg act tgg aag atc aaa ttg agg agt gca atg tac cat tca aat      528
```

```
Asn Val Thr Trp Lys Ile Lys Leu Arg Ser Ala Met Tyr His Ser Asn
                165                 170                 175

atg act cta acc atg cta gca aac ttt gta ccc ctc act ctg acc ctg      576
Met Thr Leu Thr Met Leu Ala Asn Phe Val Pro Leu Thr Leu Thr Leu
            180                 185                 190

ata tct ttt ctg ctg tta atc tgt tct ctg tgt aaa cat ctc aag aag      624
Ile Ser Phe Leu Leu Leu Ile Cys Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

atg cag ctc cat ggc aaa gga tct caa gat ccc agc acc aag gtc cac      672
Met Gln Leu His Gly Lys Gly Ser Gln Asp Pro Ser Thr Lys Val His
    210                 215                 220

ata aaa gct ttg caa act gtg acc tcc ttt ctt ctg tta tgt gcc att      720
Ile Lys Ala Leu Gln Thr Val Thr Ser Phe Leu Leu Leu Cys Ala Ile
225                 230                 235                 240

tac ttt ctg tcc atg atc ata tca gtt tgt aat ttt ggg agg ctg gaa      768
Tyr Phe Leu Ser Met Ile Ile Ser Val Cys Asn Phe Gly Arg Leu Glu
                245                 250                 255

aag caa cct gtc ttc atg ttc tgc caa gct att ata ttc agc tat cct      816
Lys Gln Pro Val Phe Met Phe Cys Gln Ala Ile Ile Phe Ser Tyr Pro
            260                 265                 270

tca acc cac cca ttc atc ctg att ttg gga aac aag aag cta aag cag      864
Ser Thr His Pro Phe Ile Leu Ile Leu Gly Asn Lys Lys Leu Lys Gln
        275                 280                 285

att ttt ctt tca gtt ttg cgg cat gtg agg tac tgg gtg aaa gac aga      912
Ile Phe Leu Ser Val Leu Arg His Val Arg Tyr Trp Val Lys Asp Arg
    290                 295                 300

agc ctt cgt ctc cat ga                                              929
Ser Leu Arg Leu His
305
```

<210> 38
<211> 309
<212> PRT
<213> Homo sapiens

<400> 38

```
Met Ile Thr Phe Leu Pro Ile Ile Phe Ser Ile Leu Ile Val Val Ile
1               5                   10                  15

Phe Val Ile Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Ser
            20                  25                  30

Ile Glu Trp Val Lys Arg Gln Lys Ile Ser Phe Val Asp Gln Ile Leu
        35                  40                  45

Thr Ala Leu Ala Val Ser Arg Val Gly Leu Leu Trp Val Leu Leu Leu
```

```
    50                  55                  60

His Trp Tyr Ala Thr Gln Leu Asn Pro Ala Phe Tyr Ser Val Glu Val
65                  70                  75                  80

Arg Ile Thr Ala Tyr Asn Val Trp Ala Val Thr Asn His Phe Ser Ser
                85                  90                  95

Trp Leu Ala Thr Ser Leu Ser Met Phe Tyr Leu Leu Arg Ile Ala Asn
            100                 105                 110

Phe Ser Asn Leu Ile Phe Leu Arg Ile Lys Arg Arg Val Lys Ser Val
        115                 120                 125

Val Leu Val Ile Leu Leu Gly Pro Leu Leu Phe Leu Val Cys His Leu
    130                 135                 140

Phe Val Ile Asn Met Asp Glu Thr Val Trp Thr Lys Glu Tyr Glu Gly
145                 150                 155                 160

Asn Val Thr Trp Lys Ile Lys Leu Arg Ser Ala Met Tyr His Ser Asn
                165                 170                 175

Met Thr Leu Thr Met Leu Ala Asn Phe Val Pro Leu Thr Leu Thr Leu
            180                 185                 190

Ile Ser Phe Leu Leu Leu Ile Cys Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

Met Gln Leu His Gly Lys Gly Ser Gln Asp Pro Ser Thr Lys Val His
    210                 215                 220

Ile Lys Ala Leu Gln Thr Val Thr Ser Phe Leu Leu Leu Cys Ala Ile
225                 230                 235                 240

Tyr Phe Leu Ser Met Ile Ile Ser Val Cys Asn Phe Gly Arg Leu Glu
                245                 250                 255

Lys Gln Pro Val Phe Met Phe Cys Gln Ala Ile Ile Phe Ser Tyr Pro
            260                 265                 270

Ser Thr His Pro Phe Ile Leu Ile Leu Gly Asn Lys Lys Leu Lys Gln
        275                 280                 285

Ile Phe Leu Ser Val Leu Arg His Val Arg Tyr Trp Val Lys Asp Arg
    290                 295                 300
```

```
Ser Leu Arg Leu His
305


<210>  39
<211>  900
<212>  DNA
<213>  Homo sapiens


<220>
<221>  CDS
<222>  (1)..(900)

<400>  39
atg atg tgt ttt ctg ctc atc att tca tca att ctg gta gtg ttt gca        48
Met Met Cys Phe Leu Leu Ile Ile Ser Ser Ile Leu Val Val Phe Ala
1               5                   10                  15

ttt gtt ctt gga aat gtt gcc aat ggc ttc ata gcy cta gta aat rtc        96
Phe Val Leu Gly Asn Val Ala Asn Gly Phe Ile Ala Leu Val Asn Xaa
            20                  25                  30

att gac tgg gtt aac aca cga aag atc tcc tca gct gag caa att ctc       144
Ile Asp Trp Val Asn Thr Arg Lys Ile Ser Ser Ala Glu Gln Ile Leu
        35                  40                  45

act gct ctg gtg gtc tcc aga att ggt tta ctc tgg gtc atg tta ttc       192
Thr Ala Leu Val Val Ser Arg Ile Gly Leu Leu Trp Val Met Leu Phe
    50                  55                  60

ctt tgg tat gca act gtg ttt aat tct gct tta tat ggt tta gaa gta       240
Leu Trp Tyr Ala Thr Val Phe Asn Ser Ala Leu Tyr Gly Leu Glu Val
65                  70                  75                  80

aga att gtt gct tct aat gcc tgg gct gta acg aac cat ttc agc atg       288
Arg Ile Val Ala Ser Asn Ala Trp Ala Val Thr Asn His Phe Ser Met
                85                  90                  95

tgg ctt gct gct agc ctc agc ata ttt tgt ttg ctc amg att gcc aat       336
Trp Leu Ala Ala Ser Leu Ser Ile Phe Cys Leu Leu Xaa Ile Ala Asn
            100                 105                 110

ttc tcc aac ctt att tct ctc cac cta aag aag aga att aag agt gtt       384
Phe Ser Asn Leu Ile Ser Leu His Leu Lys Lys Arg Ile Lys Ser Val
        115                 120                 125

gtt ctg gtg ata ctg ttg ggg ccc ttg gta ttt ytg att tgt aat ctt       432
Val Leu Val Ile Leu Leu Gly Pro Leu Val Phe Xaa Ile Cys Asn Leu
    130                 135                 140

gct gtg ata acc atg gat gag agw gtg tgg aca aaa gaa tat gaa gga       480
Ala Val Ile Thr Met Asp Glu Xaa Val Trp Thr Lys Glu Tyr Glu Gly
145                 150                 155                 160

aat gtg act tgg aag atc aaa ttg agg aat gca ata cac ctt tca agc       528
Asn Val Thr Trp Lys Ile Lys Leu Arg Asn Ala Ile His Leu Ser Ser
                165                 170                 175
```

```
ttg act gta act act cta gca aac ctc ata ccc ttt act ctg agc cta      576
Leu Thr Val Thr Thr Leu Ala Asn Leu Ile Pro Phe Thr Leu Ser Leu
            180                 185                 190

ata tgt ttt ctg ctg tta atc tgt tct ctt tgt aaa cat ctc aag aag      624
Ile Cys Phe Leu Leu Leu Ile Cys Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

atg cgg ctc cat agc aaa gga tct caa gat ccc agc acc aag gtc cat      672
Met Arg Leu His Ser Lys Gly Ser Gln Asp Pro Ser Thr Lys Val His
    210                 215                 220

rta aaa gct ttg caa act gtg acc tcc ttc ctc atg tta ttt gcc ayt      720
Xaa Lys Ala Leu Gln Thr Val Thr Ser Phe Leu Met Leu Phe Ala Xaa
225                 230                 235                 240

tac ttt ctg tgt ata atc aca tca act tgg aat ctt agg aca cag cag      768
Tyr Phe Leu Cys Ile Ile Thr Ser Thr Trp Asn Leu Arg Thr Gln Gln
                245                 250                 255

agc aaa ctt gta ctc ctg ctt tgc caa act stt gca atc atg tat cct      816
Ser Lys Leu Val Leu Leu Leu Cys Gln Thr Xaa Ala Ile Met Tyr Pro
            260                 265                 270

tca ttc cac tca ttc atc ctg att atg gga agt agg aag cta aaa cag      864
Ser Phe His Ser Phe Ile Leu Ile Met Gly Ser Arg Lys Leu Lys Gln
        275                 280                 285

acc ttt ctt tca gtt ttg tgr cag atg aca ygc tga                      900
Thr Phe Leu Ser Val Leu Xaa Gln Met Thr Xaa
    290                 295
```

<210> 40
<211> 299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (32)..(32)
<223> The 'Xaa' at location 32 stands for Val, or Ile.

<220>
<221> misc_feature
<222> (109)..(109)
<223> The 'Xaa' at location 109 stands for Lys, or Thr.

<220>
<221> misc_feature
<222> (140)..(140)
<223> The 'Xaa' at location 140 stands for Leu.

<220>
<221> misc_feature
<222> (152)..(152)

<223> The 'Xaa' at location 152 stands for Arg, or Ser.

<220>
<221> misc_feature

<222> (225)..(225)
<223> The 'Xaa' at location 225 stands for Val, or Ile.

<220>
<221> misc_feature
<222> (240)..(240)
<223> The 'Xaa' at location 240 stands for Thr, or Ile.

<220>
<221> misc_feature
<222> (267)..(267)
<223> The 'Xaa' at location 267 stands for Val, or Leu.

<220>
<221> misc_feature
<222> (295)..(295)
<223> The 'Xaa' at location 295 stands for Trp, or STOP.

<220>
<221> misc_feature
<222> (299)..(299)
<223> The 'Xaa' at location 299 stands for Arg, or Cys.

<400> 40

Met Met Cys Phe Leu Leu Ile Ile Ser Ser Ile Leu Val Val Phe Ala
1 5 10 15

Phe Val Leu Gly Asn Val Ala Asn Gly Phe Ile Ala Leu Val Asn Xaa
20 25 30

Ile Asp Trp Val Asn Thr Arg Lys Ile Ser Ser Ala Glu Gln Ile Leu
35 40 45

Thr Ala Leu Val Val Ser Arg Ile Gly Leu Leu Trp Val Met Leu Phe
50 55 60

Leu Trp Tyr Ala Thr Val Phe Asn Ser Ala Leu Tyr Gly Leu Glu Val
65 70 75 80

Arg Ile Val Ala Ser Asn Ala Trp Ala Val Thr Asn His Phe Ser Met
85 90 95

Trp Leu Ala Ala Ser Leu Ser Ile Phe Cys Leu Leu Xaa Ile Ala Asn
100 105 110

Phe Ser Asn Leu Ile Ser Leu His Leu Lys Lys Arg Ile Lys Ser Val
115 120 125

Val Leu Val Ile Leu Leu Gly Pro Leu Val Phe Xaa Ile Cys Asn Leu
130 135 140

```
Ala Val Ile Thr Met Asp Glu Xaa Val Trp Thr Lys Glu Tyr Glu Gly
145                 150                 155                 160

Asn Val Thr Trp Lys Ile Lys Leu Arg Asn Ala Ile His Leu Ser Ser
                165                 170                 175

Leu Thr Val Thr Thr Leu Ala Asn Leu Ile Pro Phe Thr Leu Ser Leu
            180                 185                 190

Ile Cys Phe Leu Leu Leu Ile Cys Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

Met Arg Leu His Ser Lys Gly Ser Gln Asp Pro Ser Thr Lys Val His
    210                 215                 220

Xaa Lys Ala Leu Gln Thr Val Thr Ser Phe Leu Met Leu Phe Ala Xaa
225                 230                 235                 240

Tyr Phe Leu Cys Ile Ile Thr Ser Thr Trp Asn Leu Arg Thr Gln Gln
                245                 250                 255

Ser Lys Leu Val Leu Leu Leu Cys Gln Thr Xaa Ala Ile Met Tyr Pro
            260                 265                 270

Ser Phe His Ser Phe Ile Leu Ile Met Gly Ser Arg Lys Leu Lys Gln
        275                 280                 285

Thr Phe Leu Ser Val Leu Xaa Gln Met Thr Xaa
    290                 295
```

<210> 41
<211> 930
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(930)

<400> 41

```
atg atg agt ttt cta cac att gtt ttt tcc att cta gta gtg gtt gca      48
Met Met Ser Phe Leu His Ile Val Phe Ser Ile Leu Val Val Val Ala
1               5                   10                  15

ttt att ctt gga aat ttt gcc aat ggc ttt ata gca ctg ata aat ttc      96
Phe Ile Leu Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Ile Asn Phe
            20                  25                  30

att gcc tgg gtc aag aga caa aag atc tcc tca gct gat caa att att     144
Ile Ala Trp Val Lys Arg Gln Lys Ile Ser Ser Ala Asp Gln Ile Ile
```

```
        35                  40                  45

gct gct ctg gcr gtc tcc aaa gtt ggt ttg ctc tgg gta ata tta tta      192
Ala Ala Leu Ala Val Ser Lys Val Gly Leu Leu Trp Val Ile Leu Leu
    50                  55                  60

cat tgg tat tca act gtg ttg aat cca act tca tct aat tta aaa gta      240
His Trp Tyr Ser Thr Val Leu Asn Pro Thr Ser Ser Asn Leu Lys Val
65                  70                  75                  80

ata att ttt att tct aat gcy tgg gca gta acc aat cat ttc agc atc      288
Ile Ile Phe Ile Ser Asn Ala Trp Ala Val Thr Asn His Phe Ser Ile
                85                  90                  95

tgg ctt gct act agc ctc agc ata ttt tat ttg ctc aag atc gtc aat      336
Trp Leu Ala Thr Ser Leu Ser Ile Phe Tyr Leu Leu Lys Ile Val Asn
            100                 105                 110

ttc tcc aga ctt att ttt cat cac tta aaa agg aag gct aag agt gta      384
Phe Ser Arg Leu Ile Phe His His Leu Lys Arg Lys Ala Lys Ser Val
        115                 120                 125

gtt ctg gtg ata gtg ttg ggg tct ttg ttc ttt ttg rtt tgt cam ctt      432
Val Leu Val Ile Val Leu Gly Ser Leu Phe Phe Leu Xaa Cys Xaa Leu
    130                 135                 140

gtg atg aaa mac acg tat ata aat gtg tgg aca gaa gaa tgt gaa gga      480
Val Met Lys Xaa Thr Tyr Ile Asn Val Trp Thr Glu Glu Cys Glu Gly
145                 150                 155                 160

aac gta act tgg aag atc aaa ctg agg aat gca atr cac ctt tcc aac      528
Asn Val Thr Trp Lys Ile Lys Leu Arg Asn Ala Xaa His Leu Ser Asn
                165                 170                 175

ttg act gta gcc atg cta gca aac ttg ata cca ttc act ctg acc ctg      576
Leu Thr Val Ala Met Leu Ala Asn Leu Ile Pro Phe Thr Leu Thr Leu
            180                 185                 190

ata tct ttt ctg ctg tta atc tac tct ctg tgt aaa cat ctg aag aag      624
Ile Ser Phe Leu Leu Leu Ile Tyr Ser Leu Cys Lys His Leu Lys Lys
        195                 200                 205

atg cag ctc cat ggc aaa gga tct caa gat ccc agc acc aag atc cac      672
Met Gln Leu His Gly Lys Gly Ser Gln Asp Pro Ser Thr Lys Ile His
    210                 215                 220

ata aaa gct ctg caa act gtg acc tcc ttc ctc rta tta ctt gcc att      720
Ile Lys Ala Leu Gln Thr Val Thr Ser Phe Leu Xaa Leu Leu Ala Ile
225                 230                 235                 240

tac ttt ctg tgt cta atc ata tcg ttt tgg aat tyt aag atg cka cca      768
Tyr Phe Leu Cys Leu Ile Ile Ser Phe Trp Asn Xaa Lys Met Xaa Pro
                245                 250                 255

aaa gaa att gtc tta atg ctt tgc caa gct ttt gga atc rta tat cca      816
Lys Glu Ile Val Leu Met Leu Cys Gln Ala Phe Gly Ile Xaa Tyr Pro
            260                 265                 270

tca ttc cac tca ttc att ctg att tgg ggg aac aag acg cta aag cag      864
Ser Phe His Ser Phe Ile Leu Ile Trp Gly Asn Lys Thr Leu Lys Gln
        275                 280                 285
```

```
acc ttt ctt tca gtt ttg tgg cag gtg act tgc tgg gca aaa gga cag      912
Thr Phe Leu Ser Val Leu Trp Gln Val Thr Cys Trp Ala Lys Gly Gln
    290                 295                 300

aac cag tca act cca tag                                          930
Asn Gln Ser Thr Pro
305
```

<210> 42
<211> 309
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (141)..(141)
<223> The 'Xaa' at location 141 stands for Val, or Ile.

<220>
<221> misc_feature
<222> (143)..(143)
<223> The 'Xaa' at location 143 stands for Gln, or His.

<220>
<221> misc_feature
<222> (148)..(148)
<223> The 'Xaa' at location 148 stands for Asn, or His.

<220>
<221> misc_feature
<222> (172)..(172)
<223> The 'Xaa' at location 172 stands for Met, or Ile.

<220>
<221> misc_feature
<222> (236)..(236)
<223> The 'Xaa' at location 236 stands for Val, or Ile.

<220>
<221> misc_feature
<222> (252)..(252)
<223> The 'Xaa' at location 252 stands for Ser, or Phe.

<220>
<221> misc_feature
<222> (255)..(255)
<223> The 'Xaa' at location 255 stands for Arg, or Leu.

<220>
<221> misc_feature
<222> (270)..(270)
<223> The 'Xaa' at location 270 stands for Val, or Ile.

<400> 42

```
Met Met Ser Phe Leu His Ile Val Phe Ser Ile Leu Val Val Val Ala
1               5                  10                  15
```

Phe Ile Leu Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Ile Asn Phe
20 25 30

Ile Ala Trp Val Lys Arg Gln Lys Ile Ser Ser Ala Asp Gln Ile Ile
35 40 45

Ala Ala Leu Ala Val Ser Lys Val Gly Leu Leu Trp Val Ile Leu Leu
50 55 60

His Trp Tyr Ser Thr Val Leu Asn Pro Thr Ser Ser Asn Leu Lys Val
65 70 75 80

Ile Ile Phe Ile Ser Asn Ala Trp Ala Val Thr Asn His Phe Ser Ile
85 90 95

Trp Leu Ala Thr Ser Leu Ser Ile Phe Tyr Leu Leu Lys Ile Val Asn
100 105 110

Phe Ser Arg Leu Ile Phe His His Leu Lys Arg Lys Ala Lys Ser Val
115 120 125

Val Leu Val Ile Val Leu Gly Ser Leu Phe Phe Leu Xaa Cys Xaa Leu
130 135 140

Val Met Lys Xaa Thr Tyr Ile Asn Val Trp Thr Glu Glu Cys Glu Gly
145 150 155 160

Asn Val Thr Trp Lys Ile Lys Leu Arg Asn Ala Xaa His Leu Ser Asn
165 170 175

Leu Thr Val Ala Met Leu Ala Asn Leu Ile Pro Phe Thr Leu Thr Leu
180 185 190

Ile Ser Phe Leu Leu Leu Ile Tyr Ser Leu Cys Lys His Leu Lys Lys
195 200 205

Met Gln Leu His Gly Lys Gly Ser Gln Asp Pro Ser Thr Lys Ile His
210 215 220

Ile Lys Ala Leu Gln Thr Val Thr Ser Phe Leu Xaa Leu Leu Ala Ile
225 230 235 240

Tyr Phe Leu Cys Leu Ile Ile Ser Phe Trp Asn Xaa Lys Met Xaa Pro
245 250 255

```
Lys Glu Ile Val Leu Met Leu Cys Gln Ala Phe Gly Ile Xaa Tyr Pro
            260                 265                 270

Ser Phe His Ser Phe Ile Leu Ile Trp Gly Asn Lys Thr Leu Lys Gln
        275                 280                 285

Thr Phe Leu Ser Val Leu Trp Gln Val Thr Cys Trp Ala Lys Gly Gln
    290                 295                 300

Asn Gln Ser Thr Pro
305
```

<210> 43
<211> 900
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(900)

<400> 43

```
atg ata act ttt cta tac att ttt ttt tca att cta ata atg gtt tta      48
Met Ile Thr Phe Leu Tyr Ile Phe Phe Ser Ile Leu Ile Met Val Leu
1               5                   10                  15

ttt gtt ctc gga aac ttt gcc aat ggc ttc ata gca ctg gta aat ttc      96
Phe Val Leu Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Phe
            20                  25                  30

att gac tgg gtg aag aga aaa aag atc tcc tca gct gac caa att ctc     144
Ile Asp Trp Val Lys Arg Lys Lys Ile Ser Ser Ala Asp Gln Ile Leu
        35                  40                  45

act gct ctg gcg gtc tcc aga att ggt ttg ctc tgg gca tta tta tta     192
Thr Ala Leu Ala Val Ser Arg Ile Gly Leu Leu Trp Ala Leu Leu Leu
    50                  55                  60

aat tgg tat tta act gtg ttg aat cca gct ttt tat agt gta gaa tta     240
Asn Trp Tyr Leu Thr Val Leu Asn Pro Ala Phe Tyr Ser Val Glu Leu
65                  70                  75                  80

aga att act tct tat aat gcc tgg gtt gta acc aac cat ttc agc atg     288
Arg Ile Thr Ser Tyr Asn Ala Trp Val Val Thr Asn His Phe Ser Met
                85                  90                  95

tgg ctt gct gct aac ctc agc ata ttt tat ttg ctc aag att gcc aat     336
Trp Leu Ala Ala Asn Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
            100                 105                 110

ttc tcc aac ctt ctt ttt ctt cat tta aag agg aga gtt agg agt gtc     384
Phe Ser Asn Leu Leu Phe Leu His Leu Lys Arg Arg Val Arg Ser Val
        115                 120                 125

att ctg gtg ata ctg ttg ggg act ttg ata ttt ttg gtt tgt cat ctt     432
Ile Leu Val Ile Leu Leu Gly Thr Leu Ile Phe Leu Val Cys His Leu
```

```
    130                 135                 140

ctt gtg gca aac atg gat gag agt atg tgg gca gaa gaa tat gaa gga      480
Leu Val Ala Asn Met Asp Glu Ser Met Trp Ala Glu Glu Tyr Glu Gly
145                 150                 155                 160

aac atg act ggg aag atg aaa ttg agg aat aca gta cat ctt tca tat      528
Asn Met Thr Gly Lys Met Lys Leu Arg Asn Thr Val His Leu Ser Tyr
                165                 170                 175

ttg act gta act acc cta tgg agc ttc ata ccc ttt act ctg tcc ctg      576
Leu Thr Val Thr Thr Leu Trp Ser Phe Ile Pro Phe Thr Leu Ser Leu
            180                 185                 190

ata tct ttt ctg atg cta atc tgt tct ctg tat aaa cat ctc aag aag      624
Ile Ser Phe Leu Met Leu Ile Cys Ser Leu Tyr Lys His Leu Lys Lys
        195                 200                 205

atg cag ctc cat gga gaa gga tcg caa gat ctc agc acc aag gtc cac      672
Met Gln Leu His Gly Glu Gly Ser Gln Asp Leu Ser Thr Lys Val His
    210                 215                 220

ata aaa gct ttg caa act ctg atc tcc ttc ctc ttg tta tgt gcc att      720
Ile Lys Ala Leu Gln Thr Leu Ile Ser Phe Leu Leu Leu Cys Ala Ile
225                 230                 235                 240

ttc ttt cta ttc cta atc gtt tcg gtt tgg agt cct agg agg ctg cgg      768
Phe Phe Leu Phe Leu Ile Val Ser Val Trp Ser Pro Arg Arg Leu Arg
                245                 250                 255

aat gac cca gtt gtc atg gtt agc aag gct gtt gga aac ata tat ctt      816
Asn Asp Pro Val Val Met Val Ser Lys Ala Val Gly Asn Ile Tyr Leu
            260                 265                 270

gca ttc gac tca ttc atc cta att tgg aga acc aag aag cta aaa cac      864
Ala Phe Asp Ser Phe Ile Leu Ile Trp Arg Thr Lys Lys Leu Lys His
        275                 280                 285

acc ttt ctt ttg att ttg tgt cag att agg tgc tga                     900
Thr Phe Leu Leu Ile Leu Cys Gln Ile Arg Cys
    290                 295
```

<210> 44
<211> 299
<212> PRT
<213> Homo sapiens

<400> 44

```
Met Ile Thr Phe Leu Tyr Ile Phe Phe Ser Ile Leu Ile Met Val Leu
1               5                   10                  15

Phe Val Leu Gly Asn Phe Ala Asn Gly Phe Ile Ala Leu Val Asn Phe
            20                  25                  30

Ile Asp Trp Val Lys Arg Lys Lys Ile Ser Ser Ala Asp Gln Ile Leu
        35                  40                  45
```

Thr Ala Leu Ala Val Ser Arg Ile Gly Leu Leu Trp Ala Leu Leu Leu
50 55 60

Asn Trp Tyr Leu Thr Val Leu Asn Pro Ala Phe Tyr Ser Val Glu Leu
65 70 75 80

Arg Ile Thr Ser Tyr Asn Ala Trp Val Val Thr Asn His Phe Ser Met
85 90 95

Trp Leu Ala Ala Asn Leu Ser Ile Phe Tyr Leu Leu Lys Ile Ala Asn
100 105 110

Phe Ser Asn Leu Leu Phe Leu His Leu Lys Arg Arg Val Arg Ser Val
115 120 125

Ile Leu Val Ile Leu Leu Gly Thr Leu Ile Phe Leu Val Cys His Leu
130 135 140

Leu Val Ala Asn Met Asp Glu Ser Met Trp Ala Glu Glu Tyr Glu Gly
145 150 155 160

Asn Met Thr Gly Lys Met Lys Leu Arg Asn Thr Val His Leu Ser Tyr
165 170 175

Leu Thr Val Thr Thr Leu Trp Ser Phe Ile Pro Phe Thr Leu Ser Leu
180 185 190

Ile Ser Phe Leu Met Leu Ile Cys Ser Leu Tyr Lys His Leu Lys Lys
195 200 205

Met Gln Leu His Gly Glu Gly Ser Gln Asp Leu Ser Thr Lys Val His
210 215 220

Ile Lys Ala Leu Gln Thr Leu Ile Ser Phe Leu Leu Leu Cys Ala Ile
225 230 235 240

Phe Phe Leu Phe Leu Ile Val Ser Val Trp Ser Pro Arg Arg Leu Arg
245 250 255

Asn Asp Pro Val Val Met Val Ser Lys Ala Val Gly Asn Ile Tyr Leu
260 265 270

Ala Phe Asp Ser Phe Ile Leu Ile Trp Arg Thr Lys Lys Leu Lys His
275 280 285

```
Thr Phe Leu Leu Ile Leu Cys Gln Ile Arg Cys
    290                 295


<210>  45
<211>  957
<212>  DNA
<213>  Homo sapiens


<220>
<221>  CDS
<222>  (1)..(957)

<400>  45
atg aat gga gac cac atg gtt cta gga tct tcg gtg act gac aag aag       48
Met Asn Gly Asp His Met Val Leu Gly Ser Ser Val Thr Asp Lys Lys
1               5                   10                  15

gcc atc atc ttg gtt acc att tta ctc ctt tta cgc ctg gta gca ata       96
Ala Ile Ile Leu Val Thr Ile Leu Leu Leu Leu Arg Leu Val Ala Ile
            20                  25                  30

gca ggc aat ggc ttc atc act gct gct ctg ggc gtg gag tgg gtg cta      144
Ala Gly Asn Gly Phe Ile Thr Ala Ala Leu Gly Val Glu Trp Val Leu
        35                  40                  45

cgg aga atg ttg ttg cct tgt gat aag tta ttg gtt agc cta ggg gcc      192
Arg Arg Met Leu Leu Pro Cys Asp Lys Leu Leu Val Ser Leu Gly Ala
    50                  55                  60

tct cgc ttc tgt ctg cag tca gtg gta atg ggt aag acc att tat gtt      240
Ser Arg Phe Cys Leu Gln Ser Val Val Met Gly Lys Thr Ile Tyr Val
65                  70                  75                  80

ttc ttg cat ccg atg gcc ttc cca tac aac cct gta ctg cag ttt cta      288
Phe Leu His Pro Met Ala Phe Pro Tyr Asn Pro Val Leu Gln Phe Leu
                85                  90                  95

gct ttc cag tgg gac ttc ctg aat gct gcc acc tta tgg tcc tct acc      336
Ala Phe Gln Trp Asp Phe Leu Asn Ala Ala Thr Leu Trp Ser Ser Thr
            100                 105                 110

tgg ctc agt gtc ttc tat tgt gtg aaa att gct acc ttc acc cac cct      384
Trp Leu Ser Val Phe Tyr Cys Val Lys Ile Ala Thr Phe Thr His Pro
        115                 120                 125

gtc ttc ttc tgg cta aag cac aag ttg tct ggg tgg cta cca tgg atg      432
Val Phe Phe Trp Leu Lys His Lys Leu Ser Gly Trp Leu Pro Trp Met
    130                 135                 140

ctc ttc agc tct gta ggg ctc tcc agc ttc acc acc att cta ttt ttc      480
Leu Phe Ser Ser Val Gly Leu Ser Ser Phe Thr Thr Ile Leu Phe Phe
145                 150                 155                 160

ata ggc aac cac aga atg tat cag aac tat tta agg aac cat cta caa      528
Ile Gly Asn His Arg Met Tyr Gln Asn Tyr Leu Arg Asn His Leu Gln
                165                 170                 175

cct tgg aat gtc act ggc gat agc ata cgg agc tac tgt gag aaa ttc      576
```

```
Pro Trp Asn Val Thr Gly Asp Ser Ile Arg Ser Tyr Cys Glu Lys Phe
            180                 185                 190

tat ctc ttc cct cta aaa wtg att act tgg aca atg ccc act gct gtc      624
Tyr Leu Phe Pro Leu Lys Xaa Ile Thr Trp Thr Met Pro Thr Ala Val
        195                 200                 205

ttt ttc att tgc atg att ttg ctc atc aca tct ctg gga aga cac agg      672
Phe Phe Ile Cys Met Ile Leu Leu Ile Thr Ser Leu Gly Arg His Arg
    210                 215                 220

aag aag gct ctc ctt aca acc tca gga ttc cga gag ccc agt gtg cag      720
Lys Lys Ala Leu Leu Thr Thr Ser Gly Phe Arg Glu Pro Ser Val Gln
225                 230                 235                 240

gca cac ata aag gct ctg ctg gct ctc ctc tct ttt gcc atg ctc ttc      768
Ala His Ile Lys Ala Leu Leu Ala Leu Leu Ser Phe Ala Met Leu Phe
                245                 250                 255

atc tca tat ttc ctg tca ctg gtg ttc agt gct gca ggt att ttt cca      816
Ile Ser Tyr Phe Leu Ser Leu Val Phe Ser Ala Ala Gly Ile Phe Pro
            260                 265                 270

cct ctg gac ttt aaa ttc tgg gtg tgg gag tca gtg att tat ctg tgt      864
Pro Leu Asp Phe Lys Phe Trp Val Trp Glu Ser Val Ile Tyr Leu Cys
        275                 280                 285

gca gca gtt cac ccc atc att ctg ctc ttc agc aac tgc agg ctg aga      912
Ala Ala Val His Pro Ile Ile Leu Leu Phe Ser Asn Cys Arg Leu Arg
    290                 295                 300

gct gtg ctg aag agt cgy cgt tcc tca agg tgt ggg aca cct tga          957
Ala Val Leu Lys Ser Arg Arg Ser Ser Arg Cys Gly Thr Pro
305                 310                 315
```

<210> 46
<211> 318
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (199)..(199)
<223> The 'Xaa' at location 199 stands for Met, or Leu.

<400> 46

```
Met Asn Gly Asp His Met Val Leu Gly Ser Ser Val Thr Asp Lys Lys
1               5                   10                  15

Ala Ile Ile Leu Val Thr Ile Leu Leu Leu Leu Arg Leu Val Ala Ile
            20                  25                  30

Ala Gly Asn Gly Phe Ile Thr Ala Ala Leu Gly Val Glu Trp Val Leu
        35                  40                  45

Arg Arg Met Leu Leu Pro Cys Asp Lys Leu Leu Val Ser Leu Gly Ala
```

50 55 60

Ser Arg Phe Cys Leu Gln Ser Val Val Met Gly Lys Thr Ile Tyr Val
65 70 75 80

Phe Leu His Pro Met Ala Phe Pro Tyr Asn Pro Val Leu Gln Phe Leu
85 90 95

Ala Phe Gln Trp Asp Phe Leu Asn Ala Ala Thr Leu Trp Ser Ser Thr
100 105 110

Trp Leu Ser Val Phe Tyr Cys Val Lys Ile Ala Thr Phe Thr His Pro
115 120 125

Val Phe Phe Trp Leu Lys His Lys Leu Ser Gly Trp Leu Pro Trp Met
130 135 140

Leu Phe Ser Ser Val Gly Leu Ser Ser Phe Thr Thr Ile Leu Phe Phe
145 150 155 160

Ile Gly Asn His Arg Met Tyr Gln Asn Tyr Leu Arg Asn His Leu Gln
165 170 175

Pro Trp Asn Val Thr Gly Asp Ser Ile Arg Ser Tyr Cys Glu Lys Phe
180 185 190

Tyr Leu Phe Pro Leu Lys Xaa Ile Thr Trp Thr Met Pro Thr Ala Val
195 200 205

Phe Phe Ile Cys Met Ile Leu Leu Ile Thr Ser Leu Gly Arg His Arg
210 215 220

Lys Lys Ala Leu Leu Thr Thr Ser Gly Phe Arg Glu Pro Ser Val Gln
225 230 235 240

Ala His Ile Lys Ala Leu Leu Ala Leu Leu Ser Phe Ala Met Leu Phe
245 250 255

Ile Ser Tyr Phe Leu Ser Leu Val Phe Ser Ala Ala Gly Ile Phe Pro
260 265 270

Pro Leu Asp Phe Lys Phe Trp Val Trp Glu Ser Val Ile Tyr Leu Cys
275 280 285

Ala Ala Val His Pro Ile Ile Leu Leu Phe Ser Asn Cys Arg Leu Arg
290 295 300

Ala Val Leu Lys Ser Arg Arg Ser Ser Arg Cys Gly Thr Pro
305 310 315

19